COT4501 Spring 2012
Homework III

This assignment has six problems and they are equally weighted. The assignment is due in class on Tuesday, February 21, 2012. There are four regular problems and two computer problems (using MATLAB). For the computer problems, turn in your results (e.g., graphs, plots, simple analysis and so on) and also a printout of your (MATLAB) code.

Problem 1

1. Show that if the vector $\mathbf{v} \neq 0$, then the matrix

$$\mathbf{H} = \mathbf{I} - 2\,\frac{\mathbf{v}\mathbf{v}^\top}{\mathbf{v}^\top\mathbf{v}}$$

   is orthogonal and symmetric.
   Solution: $\mathbf{H}$ is symmetric: $\mathbf{H}^\top = \mathbf{I}^\top - 2\frac{(\mathbf{v}\mathbf{v}^\top)^\top}{\mathbf{v}^\top\mathbf{v}} = \mathbf{I} - 2\,\frac{\mathbf{v}\mathbf{v}^\top}{\mathbf{v}^\top\mathbf{v}} = \mathbf{H}$. We have used the formula that for two matrices A, B, $(AB)^\top = B^\top A^\top$.

2. Let $\mathbf{a}$ be any nonzero vector. If $\mathbf{v} = \mathbf{a} - \alpha\,\mathbf{e}_1$, where $\alpha = \pm\,\|\mathbf{a}\|_2$, and

$$\mathbf{H} = \mathbf{I} - 2\,\frac{\mathbf{v}\mathbf{v}^\top}{\mathbf{v}^\top\mathbf{v}},$$

   show that $\mathbf{H}\mathbf{a} = \alpha\,\mathbf{e}_1$.
   Solution: By direct computation:

$$\mathbf{v}^\top\mathbf{v} = (\mathbf{a} - \alpha\,\mathbf{e}_1)^\top(\mathbf{a} - \alpha\,\mathbf{e}_1) = (\mathbf{a})^\top\mathbf{a} - 2\alpha\mathbf{a}^\top\mathbf{e}_1 + \alpha^2\mathbf{e}_1^\top\mathbf{e}_1 = 2\alpha^2 - 2\alpha\mathbf{a}^\top\mathbf{e}_1,$$

   and

$$2\mathbf{v}\mathbf{v}^\top\mathbf{a} = 2(\mathbf{a} - \alpha\mathbf{e}_1)(\mathbf{a} - \alpha\mathbf{e}_1)^\top\mathbf{a} = 2(\mathbf{a} - \alpha\mathbf{e}_1)(\alpha^2 - \alpha\mathbf{a}^\top\mathbf{e}_1).$$

   Using the two results above, we have

$$\mathbf{H}\mathbf{a} = \mathbf{a} - 2\frac{\mathbf{v}\mathbf{v}^\top\mathbf{a}}{\mathbf{v}^\top\mathbf{v}} = \mathbf{a} - \frac{2(\mathbf{a} - \alpha\mathbf{e}_1)(\alpha^2 - \alpha\mathbf{a}^\top\mathbf{e}_1)}{2\alpha^2 - 2\alpha\mathbf{a}^\top\mathbf{e}_1} = \alpha\mathbf{e}_1.$$

Problem 2 Consider the vector $\mathbf{a}$ as an $n \times 1$ matrix.

1. Write out its QR factorization, showing the matrices $\mathbf{Q}$ and $\mathbf{R}$ explicitly.
   Solution: $\mathbf{Q}$ is simply the vector (one column matrix) $\mathbf{a}/\|\mathbf{a}\|_2$, and $\mathbf{R} = \|\mathbf{a}\|_2$.

2. What is the solution to the linear least squares problem $\mathbf{a}x \simeq \mathbf{b}$, where $\mathbf{b}$ is a given n-vector.
Solution: Let $\mathbf{u} = \mathbf{a}/\|\mathbf{a}\|_2$, the $\mathbf{Q}$.. The least squares minimizes the cost function $\|\mathbf{a}x - \mathbf{b}\|_2^2 = \|\mathbf{QR}x - \mathbf{b}\|_2^2 = \|\mathbf{R}x - \mathbf{Q}^\top\mathbf{b}\|_2^2$. The solution is simply given as (using $\mathbf{Q}$ and $\mathbf{R}$ in Part 1)
$$\hat{x} = \frac{\mathbf{Q}^\top\mathbf{b}}{\mathbf{R}} = \frac{\mathbf{u}^\top\mathbf{b}}{\|\mathbf{a}\|_2} = \frac{\mathbf{a}^\top\mathbf{b}}{\|\mathbf{a}\|_2^2}.$$

3. How do you interpret the above result geometrically?
Solution: We compute the solution $\hat{x}$ (which is one single number) as the ratio between the projection of the vector $\mathbf{b}$ in the $\mathbf{a}$-direction, which is the vector $(\frac{\mathbf{a}^\top}{\|\mathbf{a}\|_2}\mathbf{b})\frac{\mathbf{a}}{\|\mathbf{a}\|} = \frac{\mathbf{a}^\top\mathbf{b}}{\|\mathbf{a}\|_2^2}\mathbf{a}$, and the vector $\mathbf{a}$. Notice that these two vectors are indeed parallel (all multiples of $\mathbf{a}$) and their ratio is the solution $\hat{x}$.

Problem 3 Consider the following matrix $\mathbf{A}$
$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ \epsilon & 0 \\ 0 & \epsilon \end{bmatrix},$$
where ϵ is a positive number smaller than $\sqrt{\epsilon_{\text{mach}}}$ in a given floating-point system. In class, we have shown that the matrix $\mathbf{A}^\top\mathbf{A}$ is singular in floating-point arithmetic. For this problem, show that if $\mathbf{A} = \mathbf{QR}$ is the reduced QR factorization for this matrix $\mathbf{A}$, then $\mathbf{R}$ is *not* singular, even in floating-point arithmetic.
Solution: Using floating-point arithmetic, the first column $\mathbf{a}_1$ of $\mathbf{A}$ has magnitude $\sqrt{1+\epsilon^2} = 1$. Therefore, the vector $\mathbf{v}_1$ for the first Householder transformation is given by $\mathbf{v}_1 = \mathbf{a}_1 + \mathbf{e}_1 = [2,\ \epsilon,\ 0]^\top$. Use this $\mathbf{H}_{\mathbf{v}_1}$, we have
$$\mathbf{H}_{\mathbf{v}_1}\mathbf{a}_1 = \begin{bmatrix} 1 \\ \epsilon \\ 0 \end{bmatrix} - \frac{2(2 = (\mathbf{v}_1^\top\mathbf{a}_1 = 2+\epsilon^2))}{4+\epsilon^2 = 4}\begin{bmatrix} 2 \\ \epsilon \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix},$$
and similarly,
$$\mathbf{H}_{\mathbf{v}_1}\mathbf{a}_2 = \begin{bmatrix} 1 \\ 0 \\ \epsilon \end{bmatrix} - \frac{2(2 = (\mathbf{v}_1^\top\mathbf{a}_2))}{4+\epsilon^2 = 4}\begin{bmatrix} 2 \\ \epsilon \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ -\epsilon \\ \epsilon \end{bmatrix}.$$
Therefore, after first Householder transformation
$$\mathbf{H}_{\mathbf{v}_1}\mathbf{A} = \begin{bmatrix} -1 & -1 \\ 0 & -\epsilon \\ 0 & \epsilon \end{bmatrix}.$$
Now for the second Householder transformation,
$$\mathbf{v}_2 = \begin{bmatrix} 0 \\ -\epsilon \\ \epsilon \end{bmatrix} - \sqrt{2}\epsilon\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -(1+\sqrt{2})\epsilon \\ \epsilon \end{bmatrix}.$$

Now we have

$$\mathbf{H}_{\mathbf{v}_2}\begin{bmatrix}-1\\-\epsilon\\\epsilon\end{bmatrix}=\begin{bmatrix}-1\\-\epsilon\\\epsilon\end{bmatrix}-\frac{2((2+\sqrt{2})\epsilon=(\mathbf{v}_2^\top\begin{bmatrix}-1\\-\epsilon\\\epsilon\end{bmatrix}))}{(4+2\sqrt{2})\epsilon^2}\begin{bmatrix}0\\-(1+\sqrt{2})\epsilon\\\epsilon\end{bmatrix}$$
$$=\begin{bmatrix}-1\\-\epsilon\\\epsilon\end{bmatrix}-\begin{bmatrix}0\\-(1+\sqrt{2})\epsilon\\\epsilon\end{bmatrix}=\begin{bmatrix}-1\\\sqrt{2}\epsilon\\0\end{bmatrix}.$$

Therefore, after the second Householder transformation, we have

$$\mathbf{H}_{\mathbf{v}_2}\mathbf{H}_{\mathbf{v}_1}\mathbf{A}=\begin{bmatrix}-1&-1\\0&\sqrt{2}\epsilon\\0&0\end{bmatrix},$$

and the $\mathbf{R}$ matrix is the non-singular matrix

$$\mathbf{R}=\begin{bmatrix}-1&-1\\0&\sqrt{2}\epsilon\end{bmatrix}.$$

Problem 4

1. What are the eigenvalues and corresponding eigenvectors of the following matrix?

$$\begin{bmatrix}1&2&-4\\0&2&1\\0&0&3\end{bmatrix}.$$

Solution: Because the matrix is upper-triangular, the three eigenvalues are $1, 2$ and 3 (the diagonal elements) with corresponding eigenvectors $[1,\ 0,\ 0]^\top$, $[2, 1,\ 0]^\top$ and $[-1,\ 1,\ 1]^\top$.

2. What are the eigenvalues of the Householder transformation

$$\mathbf{H}=\mathbf{I}-2\,\frac{\mathbf{v}\mathbf{v}^\top}{\mathbf{v}^\top\mathbf{v}},$$

where $\mathbf{v}$ is any nonzero vector?
Solution: We know that for any vector w orthogonal to v ($v^\top w=0$), we have $\mathbf{H}w=w$, and $\mathbf{H}v=-v$. There are $n-1$ (we assume that v is an n-dimensional vector) linearly independent vectors that are orthogonal to v and these are linearly independent eigenvectors of H with eigenvalue 1. v then furnishes the remaining eigenvector with eigenvalues -1. Therefore, $\mathbf{H}$ has two eigenvalues, 1 and -1. Since we know that $\mathbf{H}$ is orthogonal from Problem 1, its eigenvalues (as complex numbers) must have magnitude 1.

Computer Problem 1

1. Solve the following least squares roblem using any method you like:

$$\begin{bmatrix} 0.16 & 0.10 \\ 0.17 & 0.11 \\ 2.02 & 1.29 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \simeq \begin{bmatrix} 0.26 \\ 0.28 \\ 3.31 \end{bmatrix}.$$

2. Now solve the same least squares problem again, but this time use the slightly perturbed right-hand side,

$$\mathbf{b} = \begin{bmatrix} 0.27 \\ 0.25 \\ 3.33 \end{bmatrix}.$$

3. Compare your results from parts 1 and 2. Can you explain this difference?

Solution:

```
3.4. function cp03_04 % nearly rank deficient least squares problem
    A=[0.16 0.10; 0.17 0.11; 2.02 1.29];
    disp('(a)'); b_a=[0.26; 0.28; 3.31]; x_a = A\b_a
    disp('(b)'); b_b=[0.27; 0.25; 3.33]; x_b = A\b_b
    disp('(c)'); delx_over_x = norm(x_b-x_a)/norm(x_a)
    condA = cond(A)
    cos_theta = norm(A*x_a)/norm(b_a);
    delb_over_b = norm(b_b-b_a)/norm(b_a)
    bound = condA*cos_theta*delb_over_b
```

Computer Problem 2 A planet follows an elliptical orbit, which can be represented in a Cartesian (x, y) coordinate system by the equation

$$a\,y^2 + b\,xy + c\,x + d\,y + e = x^2.$$

- Use a library routine, or one of your own design, for linear least squares to determine the orbital parameters a, b, c, d, e given the following observations of the planet's position:

```
x: |   1.02   0.95   0.87   0.77   0.67   0.56   0.44
y: |   0.39   0.32   0.27   0.22   0.18   0.15   0.13
--------------------------------------------------------
x: |   0.30   0.16   0.01
y: |   0.12   0.13   0.15
```

In addition to printing the values for the orbital parameters, plot the resulting orbit and the given data points in the (x, y) plane.

- This least squares problem is nearly rank-deficient. To see what effect this has on the solution, use the following perturbed data

```
x: |  1.023   0.949   0.871   0.772   0.666   0.559
y: |  0.388   0.319   0.272   0.221   0.177   0.152
--------------------------------------------------
x: |  0.438   0.303   0.159   0.009
y: |  0.129   0.117   0.134   0.151
```

  and solve the least squares problem with the perturbed data. Compare the new values for the parameters with those previously computed. What effect does this difference have on the plot of the orbit? Can you explain this behavior?

```
function cp03_05 % least squares fit to planetary orbit data

x = [1.02; 0.95; 0.87; 0.77; 0.67; 0.56; 0.44; 0.30; 0.16; 0.01];
y = [0.39; 0.32; 0.27; 0.22; 0.18; 0.15; 0.13; 0.12; 0.13; 0.15];
A = [y.^2 x.*y x y ones(size(x))]; b = x.^2; disp('(a)');
figure(1); hold on;
title('Computer Problem 3.5(a) - Elliptical Orbit');
 alpha = A\b

[xs, ys] = meshgrid(-1:0.1:2, -1:0.1:2);
contour(-1:0.1:2, -1:0.1:2,...
        alpha(1)*ys.^2+alpha(2)*xs.*ys+...
        alpha(3)*xs+alpha(4)*ys+...
        alpha(5)-xs.^2, [0, 0], 'k-');
plot(x, y, 'bx'); disp('(b)');

figure(2); hold on;
title('Computer Problem 3.5(b) - Perturbed Orbit');
%x2 = x+(rand(size(x))*0.01-0.005);
% y2 = y+(rand(size(y))*0.01-0.005);
x2 = [1.023 0.949 0.871 0.772 0.666 0.559 0.438   0.303   0.159   ...
    0.009]';
y2 = [0.388 0.319 0.272 0.221 0.177 0.152 0.129   0.117   0.134   ...
    0.151]';

A2 = [y2.^2 x2.*y2 x2 y2 ones(size(x2))];
 b2 = x2.^2; alpha2 = A2\b2
[xs, ys] = meshgrid(-1:0.1:2, -1:0.1:2);
contour(-1:0.1:2, -1:0.1:2,...
         alpha2(1)*ys.^2+alpha2(2)*xs.*ys+ ...
         alpha2(3)*xs+alpha2(4)*ys+alpha2(5)-xs.^2,...
          [0, 0], 'r-');
```

```
contour(-1:0.1:2, -1:0.1:2,...
         alpha(1)*ys.^2+alpha(2)*xs.*ys+ ...
         alpha(3)*xs+alpha(4)*ys+alpha(5)-xs.^2,...
          [0, 0], 'k-');
plot(x, y, 'bx');
```