CS 422: Assignment 3

Due: Wednesday, February 27, 2008 (In Lecture)

Scoring for each problem is on a 0-to-5 scale (5 = complete success, 4 = overlooked a small detail, 3 = good start, 2 = right idea, 1 = germ of the right idea, 0 = missed the point of the problem.) One point will be deducted for insufficiently commented code. Unless otherwise stated, you are expected to utilize fully MATLAB'S vectorizing capability subject to the constraint of being flop-efficient. Test drivers and related material are posted on the course website `http://www.cs.cornell.edu/courses/cs422/2008sp/`. For each problem submit output and a listing of all scripts/functions that *you* had to write in order to produce the output. You are allowed to discuss *background* issues with other students, but the codes you submit must be your own.

P1. (Trigonometric Data Fitting)

Given $P > 0$ and a positive integer n, we wish to perform a least squares fit of the data set $(t_1, y_1), \ldots, (t_m, y_m)$ with a function of the form

$$Y_n(t) = \sum_{j=1}^{n} a_j \cos\left(\frac{2\pi j}{P} t\right) + b_j \sin\left(\frac{2\pi j}{P} t\right).$$

This just means that we want to determine $a \in \mathbb{R}^n$ and $b \in \mathbb{R}^n$ so that

$$\phi(a, b) = \sum_{k=1}^{m} \left(Y_n(t_i) - y_i\right)^2$$

is as small as possible. Refer to the optimizing a and b by $\tilde{a}^{(n)}$ and $\tilde{b}^{(n)}$. Write a MATLAB function

```
[A,B] = TrigFit(t,y,N,P)
```

that solves this problem for $n = 1{:}N$. Assume that the data set is specified by column m-vectors `t` and `y`. The output parameters `A` and `B` should be N-by-N matrices with the property that `A(1:n,n)` and `B(1:n,n)` contain $\tilde{a}^{(n)}$ and $\tilde{b}^{(n)}$ for $n = 1{:}N$. Make effective use of the MATLAB function `qr`. Submit listing and the output obtained when the test script `P1` is applied.

P2. (A QR Update Problem)

Suppose we have the QR factorization of $A \in \mathbb{R}^{m\times n}$ and that $b \in \mathbb{R}^m$, $\beta \in \mathbb{R}$, $c \in \mathbb{R}^n$ and $\theta > 0$ are given. How would you determine $x \in \mathbb{R}^n$ so that

$$\phi(x) = |c^T x - \beta|^2 \; + \; \theta \| \, Ax - b \, \|_2^2$$

is as small as possible? Write a MATLAB function

```
x = LSUpdate(Q,R,b,beta,c,theta)
```

that does this. Hint: Show that x is the solution to a structured least squares problem and make effective use of the MATLAB function `planerot`. Submit listing and the output obtained when the test script `P2` is applied.

P3. (Low Rank Least Squares)

Suppose $W \in \mathbb{R}^{m\times r}$ has rank r, $D = \text{diag}(d_1, \ldots, d_r)$, and $b \in \mathbb{R}^m$ are given. Assume that the d_i are positive. Write a MATLAB function

```
x = LowRankLS(W,d,b)
```

that returns a vector $x \in \mathbb{R}^n$ that (a) minimizes $\| \, WDW^T x - b \, \|$ and (b) has the property that at most r of its components are nonzero. Submit listing and the output obtained when the test script `P3` is applied.

P4. (Fitting Planar Data)

The set of points (x, y, z) that satisfy

$$a_1(x - x_0) + a_2(y - y_0) + a_3(z - z_0) = 0$$

defines a plane P in 3-space that passes through (x_0, y_0, z_0). The distance from a point $Q = (q_x, q_y, q_z)$ to P is given by

$$\text{dist}(Q, P) = \frac{|a_1(q_x - x_0) + a_2(q_y - y_0) + a_3(q_z - z_0)|}{\sqrt{a_1^2 + a_2^2 + a_3^2}}.$$

Given m data points $Q_1 = (x_1, y_1, z_1), \ldots, Q_m = (x_m, y_m, z_m)$ we wish to choose a_1, a_2, and a_3 so that

$$\phi(a) = \sqrt{\sum_{i=1}^{m} \text{dist}(Q_i, P)^2}$$

is minimum assuming that the centroid[1] of the data set is on P. Write a MATLAB function

```
[a,phiMin] = BestPlane(x,y,z)
```

that does this. The input parameters should be column m-vectors that define the data set. The output parameter `a` should be a column 3-vector that defines the best plane and `phiMin` should be the associated value of ϕ. Submit listing and the output obtained when the test script `P4` is applied.

[1] The xyz coordinates of the data set centroid are obtained by averaging the x, y, and z coordinates of the data.