CS 6210 Assignment 4   Due: 10/23/15 (Fri) at 11pm

Scoring for each problem is on a 0-to-5 scale (5 = complete success, 4 = overlooked a small detail, 3 = good job on half the problem, 2 = OK job on half the problem, 1 = germ of a relevant solution idea, 0 = missed the point of the problem.) Independent of this, one point will be deducted for insufficiently commented code. Test code and related material are posted on the course website `http://www.cs.cornell.edu/courses/cs6210/2015fa/`. All solution M-Files must be submitted through the CMS system. You are allowed to discuss *background* issues with other students, but the codes you submit must be your own. If part of your solution to a problem is based on something found on the Web or in the published literature, then include a citation comment.

Topics: QR Factorization, SVD, Rank Degeneracy, Least Squares Fitting

1 Low Rank Least Squares

Consider the least squares problem

$$\min_{x \in \mathbb{R}^n} \quad \| FG^T x - b \|_2$$

where $F \in \mathbb{R}^{m\times r}$, $G \in \mathbb{R}^{n\times r}$, and $b \in \mathbb{R}^m$. Assume that $\text{rank}(F) = \text{rank}(G) = r < n \leq m$. This is a rank deficient LS problem. Assume that we want to compute x_{LS}, the minimum 2-norm solution. Note that if

$$U^T(FG^T)V = \Sigma = \text{diag}(\sigma_1, \ldots, \sigma_n)$$

is the singular value decomposition, then

$$x_{LS} = \sum_{j=1}^{r} \frac{U(:,j)^T b}{\sigma_j} V(:,j)$$

Here is a MATLAB function that returns x_{LS} using this expression:

```
  function xLS = LowRankLS(F,G,b)
% F is m-by-r and rank(F) = r
% G is n-by-r and rank(G) = r
% b is m-by-1
% xLS is the minimum norm solution to min ||(FG')x - b ||.

[m,r] = size(F);
M = F*G';
[U,S,V] = svd(M,0);
xLS = V(:,1:r)*((U(:,1:r)'*b)./diag(S(1:r,1:r)));
```

This implementation involves $O(mn)$ storage and $O(mn^2)$ work. Develop an improved implementation that involves just $O(mr)$ storage and $O(mr^2)$ work. (This is a big deal if $r << n$.) Hints. Avoid forming FG^T and make effective use of the MATLAB function `qr`. Think hard about the nullspace of FG^T. This subspace is important because x_{LS} is in its orthogonal complement. Submit your implementation of `LowRankLS` to CMS.

2 QR With RHS-Driven Column Pivoting

This problem is about solving the rank-deficient LS problem using QR with column pivoting so start by reviewing §5.4.2 in GVL4. Recall that if $A \in \mathbb{R}^{m\times n}$ and $\text{rank}(A) = r \leq n \leq m$ and

$$Q^T A \Pi = \begin{bmatrix} R_{11} & R_{12} \\ 0 & 0 \end{bmatrix}$$

where $Q \in \mathbb{R}^{m \times m}$ is orthogonal and $R_{11} \in \mathbb{R}^{r \times r}$ is upper triangular and nonsingular, then the basic solution to the min $\| Ax - b \|$ problem is given by

$$x_B \;=\; \Pi \left[\begin{array}{c} R_{11}^{-1} Q(:,1{:}r)^T b \\ 0 \end{array} \right]$$

In practice, all matrices have full column rank so we normally compute

$$Q^T A \Pi \;=\; \left[\begin{array}{c} R_{11} \\ 0 \end{array} \right] \qquad R_{11} \in \mathbb{R}^{n \times n}$$

and then decide upon rank by looking at the diagonal elements of R_{11}. When the column pivoting is based on 2-norms (as in GVL4 Algorithnm 5.4.1) we obtain obtain the following procedure with the help of the Householder vector function `house`:

$P = I_n$

$\tilde{b} = b$

for $j = 1{:}n$

  Determine $i \geq j$ so that $\| A(j{:}m, i) \|_2^2 = \max\{\| A(j{:}m, k) \|_2^2\}_{k=j}^n$

  Let $P_j \in \mathbb{R}^{n \times n}$ be the identity I_n with columns i and j swapped.

  $A = AP_j$

  $P = PP_j$

  $[v, \beta] = \texttt{house}(A(j{:}m, j))$

  $A(j{:}m, j{:}n) = (I - \beta vv^T)A(j{:}m, j{:}n)$

  $\tilde{b}(j{:}m) = (I - \beta vv^T)\tilde{b}(j{:}m)$

end

In step j we swap in that column that is "maximally different" from all the previously chosen columns. The algorithm does this by looking for the column in $A(j{:}m, j{:}n)$ that has the largest 2-norm. We will call this *2-norm column pivoting.* Note that x_B can be computed from the output of this procedure.

An alternative pivot strategy results if at the start of each step we "swap in" the column that maximizes the reduction of the residual. This is accomplished if we adopt the following pivot strategy[1]:

$$\text{Determine } i \geq j \text{ so that } |\cos(b(j{:}m)^T A(j{:}m, i))| = \max\{|\cos(b(j{:}m)^T A(j{:}m, k))|\}_{k=j}^n$$

We will call this pivot strategy the *rhs column pivoting strategy.* Implement the following function so that it performs as specified:

```
  function [R1,P,btilde] = BasicLSrhs(A,b)
% A is m-by-n with m>=n and b is m-by-1
% Computes an orthogonal Q and a permutation P so that Q'*A*P = R is
% upper triangular and P is a permutation determined by the rhs
% column pivoting strategy.
% btilde = Q'*b and R1 = R(1:n,1:n).
```

Structure your implementation after the pseudocode given above. Make effective use of the function `house` that is available on the course website. Make sure your updates of A, b, and P are efficient. Pay attention to the computation of the cosines so that you do not introduce an $O(mn^2)$ overhead. In particular, develop an efficient way to update the cosines analogous to how the column norms are updated when we do QR with 2-norm column pivoting. You can check your code by comparing `xLS` and `xB` in

```
xLS = A\b;
[R1,P,btilde] = BasicLSrhs(A,b);
xB = P*(R1\btilde(1:n));
```

Submit `BasicLSrhs` to CMS.

[1]The cosine between two vectors u and v is given by $\cos(u, v) = u^T v/(\| u \| \| v \|)$.

3 A Structured Rank-Reduction Perturbation

Suppose $A \in \mathbb{R}^{m\times(p+s)}$ with $m \geq p+s$ and that

$$A \;=\; [A_1 \mid A_2\,] \qquad A_1 \in \mathbb{R}^{m\times p},\; A_2 \in \mathbb{R}^{m\times s}$$

Assume that A has full column rank. What is the minimum value of $\parallel E_2 \parallel_2$ subject to the constraint that

$$\tilde{A} \;=\; [A_1 \mid A_2 + E_2\,] \qquad E_2 \in \mathbb{R}^{m\times s}$$

is rank deficient? In other words, we want to make A rank deficient by minimally perturbing its last s columns.

Hints. (1) If $s = n$ then this is an easy SVD problem: $E_2^{(opt)} = \sigma_n u_n v_n^T$. See GVL4 §2.4. (2) Use the QR factorization to convert the given problem into a square upper triangular problem and then determine the "best" E_{12} and E_{22} so that

$$\left[\begin{array}{cc} R_{11} & R_{12}+E_{12} \\ 0 & R_{22}+E_{22} \end{array}\right]. \tag{1}$$

is singular. (3) If the matrix in (1) is singular, then what can you say about $R_{22}+E_{22}$? Complete the following function so that it performs as specified:

```
  function [normE,E] = RightHalfPerturb(A,s)
% A is mxn with full column rank and 1<=s<n
% E is an m-by-s matrix with the property that [ A(:,1:n-s) A(:,n-s+1:n+e)] is
% rank deficient and no other matrix with this property has small 2-norm.
% normE is the 2-norm of E
```

Submit your implementation to CMS.

4 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 (x_0, y_0, z_0) is the the centroid of the data set[2]. Write a MATLAB function

```
[a,Q0,phiMin] = BestPlane(Q)
```

that does this. The input matrix Q is m-by-3 with rows that specify the points that are to be fit with a plane. The output vector a should be a row-oriented 3-vector that specifies the normal associated with the fitting plane. `Q0` should be the centroid of the data set (a length-3 row vector) and `phiMin` should be the associated optimal value of ϕ. Submit `BestPlane` to CMS.

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