CS 621: Final Exam Solution

Monday, December 10, 2007

9:00-11:30

Name: ____________________________________

Problem 1	15 points	
Problem 2	10 points	
Problem 3	15 points	
Problem 4	15 points	
Problem 5	30 points	
Problem 6	15 points	
	100 points	

1. (a) (8 points) A 2-by-2 linear system $Ax = b$ is solved via Gaussian elimination with partial pivoting on a computer with unit roundoff 10^{-16}. Here is the computed result:

$$\hat{x} = \begin{bmatrix} 123.45678912345678912 \\ 0.00012345678912345 \end{bmatrix}$$

Underline the digits that are most likely correct given that $\kappa_\infty(A) \approx 10^6$. Explain.

Solution:

The basic result about relative error (in the infinity norm):

$$\| \hat{x} - x \| / \| x \| \approx \text{unit roundoff} \times \kappa(A) \approx 10^{-10}$$

Since $\| x \| \approx 10^2$ we have $\| \hat{x} - x \| \approx 10^{-8}$. So in each component we are justified in underlining digits through the 7th place to the left of the decimal point.

$$\hat{x} = \begin{bmatrix} \texttt{123.4567891**********} \\ \texttt{000.0001234**********} \end{bmatrix}$$

Off by one is fine.

Four points for saying relative error 10^{-10} and then 2 points for translating what that means for the first component and 2 points for what that means for the second component.

(b) (7 points) It is often said that when Gaussian elimination with partial pivoting is applied to A, then the condition of U is roughly the condition of A. Can you make this more precise if we know that the condition of L is less than 10?

Solution:

Since

$$PA = LU \Rightarrow \| A \| \leq \| L \| \| U \|$$

and

$$A^{-1}P^T = U^{-1}L^{-1} \Rightarrow \| A^{-1} \| \leq \| L^{-1} \| \| U^{-1} \|$$

we have $\kappa(A) \leq \kappa(L)\kappa(U) = 10\ \kappa(U)$. (Five points for this.) Likewise,

$$L^{-1}PA = U \Rightarrow \| U \| \leq \| L^{-1} \| \| A \|$$

and

$$U^{-1} = A^{-1}P^T L \Rightarrow \| U^{-1} \| \leq \| A^{-1} \| \| L \|$$

and so $\kappa(U) \leq \kappa(L)\kappa A$. (2 points for this.) Thus, $\kappa(A)/10 \leq \kappa(U) \leq 10\kappa(A)$.

2. (10 points) We say that a matrix $A \in \mathbb{R}^{2n \times 2n}$ is *symplectic* if

$$J^T A^T J = A^{-1}$$

where

$$J = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}.$$

Complete the following function so that it performs as specified:

```
  function Q = SympProd(U,V)
% Q = U*V where U and V are 2n-by-2n orthogononal symplectic matrices
```

Solution:

We check to see what $J^T A^T J = A^{-1}$ means when

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

is orthogonal. Since $A^T = A^{-1}$, we equate blocks in $J^T A^T J = A^T$:

$$\begin{bmatrix} A_{22}^T & -A_{12}^T \\ -A_{21}^T & A_{11}^T \end{bmatrix} = \begin{bmatrix} A_{11}^T & A_{21}^T \\ A_{12}^T & A_{22}^T \end{bmatrix}$$

Thus, an orthogonal symplectic Q has the form

$$Q = \begin{bmatrix} Q_1 & Q_2 \\ -Q_2 & Q_1 \end{bmatrix}$$

If U and V are orthogonal symplectic then

$$J^T (UV)^T J = (J^T V^T J)(J^T U^T J) = V^T U^T = (UV)^T$$

so UV is symplectic. Thus,

```
[m,m] = size(U); n = m/2; alfa = 1:n; beta=n+1:m;
U1 = U(alfa,alfa); U2 = U(alfa,beta);
V1 = V(alfa,alfa); V2 = V(alfa,beta);
Q1 = U1*V1 - U2*V2; Q2 = U1*V2 + U2*V1;
Q = [ Q1 Q2 ; -Q2 Q1]
```

Minor deductions of 1, 2 points for minor errors.

3. (15 points) Complete the following function so that it performs as specified

```
  function delta = Range(A,i,j)
% A is n-by-n symmetric positive definite and 1 <= j < i <= n
% delta is the smallest possible real number (in absolute value)
% such that A becomes singular when A(i,j) is replaced by
% A(i,j) + delta and A(j,i) is replaced by A(j,i) = delta
```

Make effective use of the Sherman-Morrison-Woodbury formula

$$(A + UV^T)^{-1} = A^{-1} - A^{-1}UMV^TA^{-1} \qquad M = I + V^TA^{-1}U$$

Solution:

Note that by changing these entries we are making a rank-2 change to A:

$$A + \delta(e_ie_j^T + e_je_i^T) = A + UV^T$$

where $U = [e_i\ e_j]$ and $V = [\delta e_j\ \delta e_i]$. (5 points for this)

The point at which $A + UV^T$ "goes singular" is the point at which the 2-by-2 matrix $I + V^TA^{-1}U$ "goes singular" for then the inverse of $A + UV^T$ does not exist. If $X = A^{-1}$ then

$$M = I + V^TA^{-1}U = I + \delta[e_j\ e_i]^TX[e_i\ e_j] \ = \left[\begin{array}{cc} 1 + \delta x_{ji} & \delta x_{jj} \\ \delta x_{ii} & 1 + \delta x_{ij} \end{array}\right]$$

(5 points for the relevance of M and computing it as a function of δ.)

Looking at the determinant of this, we want δ to be the smaller root of the quadratic

$$(1 + \delta x_{ij})^2 - \delta^2 x_{ii}x_{jj} = \delta^2(x_{ij}^2 - x_{ii}x_{jj}) + 2x_{ij}\delta + 1 = 0$$

$$\delta = \frac{-2x_{ij} \pm \sqrt{4x_{ij}^2 - 4(x_{ij}^2 - x_{ii}x_{jj})}}{2(x_{ij}^2 - x_{ii}x_{jj})} \ = \frac{-x_{ij} \pm \sqrt{x_{ii}x_{jj}}}{x_{ij}^2 - x_{ii}x_{jj}}$$

(5 points for setting up the quadratic and specifying the smaller root.)

```
n = length(A);
G = chol(A)';
ei = zeros(n,1); ei(i) = 1; y = G'\(G \ ei);
ej = zeros(n,1); ej(j) = 1; z = G'\(G \ ej);
xii = y(i);
xij = y(j);
xjj = z(j)
if xij<0
   delta = (-xij - sqrt(xii*xjj))/(xij^2 - xii*xjj);
else
   delta = (-xij + sqrt(xii*xjj))/(xij^2 - xii*xjj);
end
```

4. (15 points) This problem is about solving linear systems of the form $(A^2 - \lambda^2 I)x = b$. Complete the following function so that it performs as specified and is efficient:

```
   function X = MultShiftSolver(A,b,d)
%  A is n-by-n, b is n-by-1, d is m-by-1 with m>>n
%  For k=1:m, X(:,k) solves the linear system (A^2 - d(k)^2 I)z = b.
%  Assume that each linear system is nonsingular.
```

You are allowed to use the "\" operator and may assume that it fully exploits bandedness.

Solution:

$$(A^2 - d_k^2 I)z = (A + d_k I)(A - d_k I)z = b$$

If $A = UHU^T$ is the Hessenberg factorization of A then this liner system transforms to

$$(H - d_k I)(H + d_k I)z = c$$

where $x = Uz$ and $c = U^T b$. This system can be solved in $O(n^2)$ flops since it is a pair of Hessenberg system solves. Thus we obtain

```
[n,n] = size(A);
[U,H] = hess(A);
c = U'*b;
m = length(d);
X = zeros(n,m);
for k=1:m
   H1 = H - d(k)*eye(n,n);
   H2 = H + d(k)*eye(n,n);
   X(:,k) = U*(H2\(H1\c));
end
```

No deductions for working with the schur form or for computing the hess or schur form of A^2. Indeed, for appropriate choices of m and n these might be smart things to do.

Correct solutions that involve computing an LU factorization for each d_k received about 5 points. Incorrect solutions that tried to live with a single LU or QR factorization also received around 5 points depending on what was said and shown.

5(a). (8 points) Suppose $A \in \mathbb{R}^{m_1 \times n}$ and $B \in \mathbb{R}^{m_2 \times n}$ and let S_A and S_B denote their null spaces. How would you use the SVD to compute a basis for the the intersection of S_A and S_B? Justify your answer.

Solution: Let

$$C = \begin{bmatrix} A \\ B \end{bmatrix}$$

It is easy to see that the null space of this matrix is the intersection of S_A and S_B. So if $U^T C V = \Sigma$ and $\sigma_{r+1}, \ldots, \sigma_n$ are the small singular values, then the columns of $V(:, r+1{:}n))$ are the required basis vectors.

Half credit for just computing the null spaces of A and B. It is possible to subsequently work with these subspaces and get the right answer, credit given accordingly.

(5b). (7 points) Suppose $A \in \mathbb{R}^{n \times n}$ is symmetric with distinct eigenvalues. Show how to compute a rank-2 matrix E so that $A + E$ has a repeated eigenvalue and $\| E \|_2 = \delta/2$ where δ is the smallest gap between any pair of eigenvalues.

Solution: Suppose $Q^T A Q = D = \text{diag}(d_1, \ldots, d_n)$ is the Schur decomposition of A:

$$A = \sum_{i=1}^{n} d_i q_i q_i^T$$

where $Q = [q_1, \ldots, q_n]$ Without loss of generality we may assume that the gap between d_1 and d_2 is the minimum gap. The matrix

$$\tilde{A} = \frac{d_1 + d_2}{2} q_1 q_1^T + \frac{d_1 + d_2}{2} q_2 q_2^T + \sum_{i=3}^{n} d_i q_i q_i^T$$

has a repeated eigenvalue and

$$\tilde{A} - A = \frac{d_2 - d_1}{2} q_1 q_1^T + \frac{d_1 - d_2}{2} q_2 q_2^T = E$$

has 2-norm equal to $|d_1 - d_2|/2$. Minus 2 or 3 if δ is right but the rank-2 is incorrectly specified.

(5c). (8 points) A large sparse real orthogonal matrix Q has only one eigenvalue in the right half plane. How would you compute the associated eigenvector using the power method? Justify your answer.

Solution:

All the eigenvalues of Q are on the unit circle. If Q has only one eigenvalue in the right half plane then it must be real and it is therefore equal to one. Since all the eigenvalues are on the unit circle,the power method would not go anywhere. But if we apply the power method to $Q+I$ then the dominant eigenvalue is 2 and the associated eigenvector is what we want. The ratio of the second largest eigenvalue to this is $\sqrt{2}/2$.

Three, four, five points if you tried to work with a modification like $(Q-\mu I)^{-1}$ or $(Q+Q^T)/2$

(5d). (7 points) Suppose A is symmetric and positive definite and that $b \in \mathbb{R}^n$. After k steps of the Lanczos process we have the factorization $AQ_k = Q_kT_k + r_ke_k^T$ where $Q_k \in \mathbb{R}^{n\times k}$ has orthonormal columns, $T_k \in \mathbb{R}^{k\times k}$ is tridiagonal, $r_k \in \mathbb{R}^n$ satisfies $Q_k^Tr_k = 0$, and $e_k = I_k(:,k)$. Assume that Q_k is easily accessed and that $Q_k(:,1) = b$. How would you minimize

$$\phi(x) = \frac{1}{2}x^TAx - x^Tb$$

subject to the constraint that $x \in \text{span}\{b, Ab, \ldots, A^{k-1}b\}$?

Solution: Since the columns of Q_k are a basis for $\text{span}\{b, Ab, \ldots, A^{k-1}b\}$, we are looking for a vector y that minimizes

$$\phi(Q_ky) = \frac{1}{2}(Q_ky)^TA(Q_ky) - (Q_ky)^Tb = \frac{1}{2}y^TT_ky - y^TQ_k^b$$

The optimum y_{opt} solves the symmetric, positive definite tridiagonal system $T_ky = Q_k^Tb$. Thus, $x_{0pt} = Q_ky_{opt}$.

6. (15 points) Assume that $A \in \mathbb{R}^{m\times m}$ and $C \in \mathbb{R}^{n\times n}$ are upper triangular. Suppose $B \in \mathbb{R}^{m\times n}$. By comparing columns in the equation $X + AXC = B$, derive an algorithm for computing $X \in \mathbb{R}^{m\times n}$ assuming that it exists. Informal `Matlab` pseudocode is sufficient. Under what conditions might a solution fail to exist?

Solution: Comparing column k:

$$X(:,k) + \sum_{i=1}^{k} A(c_{ik}X(:,i)) = B(:,k)$$

Five points for this.

Thus,

$$(I + c_{kk}A)X(:,k) = B(:,k) - A\left(\sum_{i=1}^{k-1} c_{ik}X(:,i)\right)$$

Thus, we can resolve X by using this formula for $k = 1, 2, \ldots, n$. Seven points for this. (Minus 1 if A is inside the summation.)

The m-by-m triangular matrix $(I + c_{kk}A)$ has to be nonsingular. Thus, $(1 + c_{kk}a_{jj}) \neq 0$. In other words, if λ is an eigenvalue of A then $-1/\lambda$ had better not be an eigenvalue of C. Three points for this.

Some MATLAB Functions

LU Factorization

```
[L,U,P] = LU(X) returns unit lower triangular matrix L, upper
triangular matrix U, and permutation matrix P so that
P*X = L*U.
```

Cholesky Factorization

```
R = CHOL(X)  returns an upper triangular R so that R'*R = X
where X is symmetric and positive definite.
```

QR Factorization

```
[Q,R,E] = QR(A) produces unitary Q, upper triangular R and a
permutation matrix E so that A*E = Q*R. The column permutation E is
chosen so that ABS(DIAG(R)) is decreasing.
```

Note: `QR(X,0)` produces the thin QR factorization.

Singular Value Decomposition

```
[U,S,V] = SVD(X) produces a diagonal matrix S, of the same
dimension as X and with nonnegative diagonal elements in
decreasing order, and unitary matrices U and V so that
X = U*S*V'.
```

Note: `SVD(X,0)` produces the thin SVD.

Hessenberg Decomposition

```
HESS   Hessenberg form.
   H = HESS(A) is the Hessenberg form of the matrix A.
   The Hessenberg form of a matrix is zero below the first
   subdiagonal and has the same eigenvalues as A. If the matrix
   is symmetric or Hermitian, the form is tridiagonal.
```

Schur Decomposition

```
SCHUR  Schur decomposition.
   [U,T] = SCHUR(X) produces a quasitriangular Schur matrix T and
   an orthogonal matrix U so that X = U*T*U' and U'*U = EYE(SIZE(U)).
   X must be square.
```