CS 4220: Prelim Solutions

1.	25 points	20
2.	25 points	15
3.	25 points	23
4.	25 points	10 + 6
	100 points	74

Median = 78. Grade **Guidelines**: A(85-100), B(65-80), C(50-60).

1. Complete the following function so that it performs as specified. Your solution should be vectorized and flop efficient. You may use \.

```
  function  X = BigSol(L,U,B)
% L is n-by-n, nonsingular, and lower triangular.
% U is n-by-n, nonsingular, and upper triangular.
% B is n-by-n and lower triangular.
% X is n-by-n and solves L*U*X = B.
```

Note that the solution to $LY = B$ is lower triangular. This means that

$$Y(:,k) \;=\; \left[\begin{array}{c} \text{zeros(k-1,1)} \\ L(k{:}n,k{:}n)\backslash B(k{:}n,k) \end{array}\right]$$

Thus, the full credit solution is

```
[n,n] = size(L);
X = zeros(n,n);
for k=1:n
   y(k:n) = L(k:n,k:n)\B(k:n,k);
   X(:,k) = U\y;
end
```

-7 for $Y = U\backslash(L\backslash B)$. -3 if you exploit structure for $LY = B$ but erroneously try to exploit structure in $UX = Y$. (The X matrix is full–there is no structure.) Up to 2 points for scribble that shows you were trying to find structure.

2. Suppose $A \in \mathbb{R}^{n\times n}$ has singular values $\sigma_1 \geq \cdots \geq \sigma_n > 0$ and that $b \in \mathbb{R}^n$. It can be shown that if $E \in \mathbb{R}^{n\times n}$ satisfies

$$\| E \|_2 \;\leq\; \frac{1}{2}\sigma_n$$

then $A+E$ is nonsingular and the vectors $x = A^{-1}b$ and $\hat{x} = (A+E)^{-1}b$ satisfy

$$\frac{\| \hat{x} - x \|_2}{\| x \|_2} \;\leq\; 2\frac{\| E \|_2}{\sigma_n}.$$

Use this result to explain and make more precise the following statement:

> If A is not too badly conditioned with respect to the unit roundoff u, then the computed solution to $Ax = b$ obtained via Gaussian elimination with pivoting will typically have relative error that is approximately the product of u and the condition of A.

15 points for the condition of A:

$$\kappa_2(A) \;=\; \| A \|_2 \| A^{-1} \|_2 \;=\; \frac{\sigma_1}{\sigma_n}$$

Up to 10 points for reasoning as follows... Thus

$$\frac{\| \hat{x} - x \|_2}{\| x \|_2} \;\leq\; 2\frac{\| E \|_2}{\| A \|_2}\kappa_2(A). \qquad (1)$$

and

$$\frac{\| E \|_2}{\| A \|_2}\kappa_2(A) \;\leq\; \frac{1}{2} \qquad (2)$$

Gaussian elimination with pivoting produces an $\hat{x}$ that satisfies $(A+E)\hat{x} \approx b$ where $\| E \|_2 \approx u\| A \|_2$. Using (2) we conclude that if $u\kappa_2$ is smaller than $O(1)$ then a nearby system will be solved and

$$\frac{\| \hat{x} - x \|_2}{\| x \|_2} \;\approx\; u\kappa_2(A)$$

3(a). Suppose

$$A \;=\; \begin{bmatrix} d_1 & 0 & 0 & e_1 \\ 0 & d_2 & 0 & e_2 \\ 0 & 0 & d_3 & e_3 \\ e_1 & e_2 & e_3 & d_4 \end{bmatrix}$$

is positive definite. Show how to compute

$$G \;=\; \begin{bmatrix} g_1 & 0 & 0 & 0 \\ 0 & g_2 & 0 & 0 \\ 0 & 0 & g_3 & 0 \\ h_1 & h_2 & h_3 & g_4 \end{bmatrix}$$

so that $A = GG^T$.

$d_1 = g_1^2 \;\Rightarrow\; g_1 = \sqrt{d_1}$
$d_2 = g_2^2 \;\Rightarrow\; g_2 = \sqrt{d_2}$
$d_3 = g_3^2 \;\Rightarrow\; g_3 = \sqrt{d_3}$
$e_1 = h_1 g_1 \;\Rightarrow\; h_1 = e_1/g_1$
$e_2 = h_2 g_2 \;\Rightarrow\; h_2 = e_2/g_2$
$e_3 = h_3 g_3 \;\Rightarrow\; h_3 = e_3/g_3$
$d_4 = e_1^2 + e_2^2 + e_3^2 + g_4^2 \;\Rightarrow\; g_4 = \sqrt{d_4 - e_1^2 + e_2^2 + e_3^2}$

Worth 12 points. -3 if g_4 recipe wrong.

3(b). Complete the following function so that it performs as specified. Your solution should be vectorized and flop-efficient.

```
  function [g,h] = Arrow(d,e)
% Assume that A is a symmetric positive definite matrix that is zero everywhere
% except on its diagonal, last column, and last row.
% d is a column n-vector that is the diagonal of A and e = A(1:n-1,n).
% Let G be a lower triangular matrix so that A = GG'.
% g is a column n-vector that is the diagonal of G and h = G(n,1:n-1)'.

  n = length(d);
  g(1:n-1) = sqrt(d(1:n-1));
  h = e./g(1:n-1);
  g(n) = sqrt(d(n)-e'*e)
```

Worth 13 points. -2 if not vectorized. -2 if $g(n)$ computation wrong.

4(a). Complete the following function so that it performs as specified:

```
  function [B,b] = Update(A,a,v)
% A is m-by-n matrix.
% a is a 1-by-n vector with a(i) = A(:,i)'*A(:,i), i=1:n.
% v is m-by-1 and has unit 2-norm.
% B = H*A where H = I - 2*v*v'.
% b is a 1-by-n vector with b(i) = B(2:m,i)'*B(2:m,i), i=1:n.
```

Note that $B = (I - 2vv^T)A = A - 2vv^TA = A - (2v)(v^TA)$. For the norm part, if Q is orthogonal and

$$Qx \;=\; \begin{bmatrix} \alpha \\ w \end{bmatrix}$$

where α is the top component, then because 2-norms are preserved

$$x^Tx \;=\; \alpha^2 + w^Tw$$

So you basically have to apply the formula $w^Tw = x^Tx - \alpha^2$ to each column, i.e., $B(2{:}m, i)^TB(2{:}m, i) = A(:, i)^TA(:, i) - B(1, i)^2$.

```
B = A - (2*v)*(v'*A);                 10 points
b = a - B(1,1:n).^2                    5 points
```

4(b). If $C \in \mathbb{R}^{m\times n}$ then the Frobenius norm is given by

$$\| \, C \, \|_F \;=\; \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}^2}$$

Given $B \in \mathbb{R}^{m\times m}$, how would you compute a matrix $X \in \mathbb{R}^{n\times n}$ that minimizes $\| \, DXD^T - B \, \|_F$ where $D \in \mathbb{R}^{m\times n}$ is diagonal with $m \geq n$. Make no assumptions about $d_{11}, \ldots, d_{nn}$.

Take a look at the 3-by-2 instance of the problem:

$$DXD^T - B \;=\; \begin{bmatrix} x_{11}d_1d_1 & x_{12}d_1d_2 & 0 \\ x_{21}d_2d_1 & x_{22}d_2d_2 & 0 \\ 0 & 0 & 0 \end{bmatrix} - \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}$$

To minimize the sum of squares we set $x_{11} = b_{11}/d_{11}^2$, $x_{12} = b_{12}/(d_{11}d_{22})$, $x_{21} = b_{21}/(d_{11}d_{22})$, and $x_{22}/(d_{22}^2$. If any of the denominators is zero then the value assigned to the corresponding x_{ij} has no effect on the sum of squares. (Up to 8 points for the right idea from a small example like this.)

10-point solution: x_{ij} can take on any value unless d_{ii} and d_{jj} are nonzero, in which case we must set $x_{ij} = b_{ij}/(d_{ii}d_{jj})$.