Math 170A
Final Exam

Name: Answer Key
Student ID number: ____________________________________________

Instructions: Answers without work may be given no credit at the grader's discretion. The test is out of 75 points.

This cover page may be used at scratch paper. However, *all* final work must be on the page with the related question. Do not remove this sheet.
1. Explain why finding and using the LU decomposition is preferable to finding and using $A^{-1}$. (3 pts)

It takes fewer flops ($\approx \frac{2}{3} n^3$ instead of $2 n^3$).

or

The LU decomposition can take advantage of sparse matrices, but $A^{-1}$ will usually not be sparse, even if $A$ is.

2. Explain why finding and using the LU decomposition is preferable to just directly solving $Ax = b$ through any kind of row reduction. (3 pts)

You can use the LU decomposition to solve $Ax = b$ for several different $b$'s very cheaply ($O(n^2)$ each). But if you just use row reduction, each new $b$ takes $O(n^3)$. 
3. Explain why the Cholesky decomposition is usually preferable to the LU decomposition, if the matrix is positive definite. (3 pts)

It takes about half the flops as LU \( \left( \frac{n^3}{3} \text{ vs } \frac{2}{3}n^3 \right) \).

Or

It is more numerically stable. For instance, it is always backwards stable, while the LU decomp may not be.

4. What is catastrophic cancellation? (3 pts)

Catastrophic cancellation is when you add (or subtract) two very similar floating points, and they almost exactly cancel out. This causes you to rapidly lose several digits of accuracy. For example, using 4 digit floating points

\[ 3.196 - 3.195 = 1.000 \times 10^{-3}. \]

But on the left, we presumably have 4 digits of accuracy, but on the right, only the first digit has any real meaning or accuracy.
5. Calculate $\|A\|_2$ if $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. (3 pts)

$A^TA = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$. Eigenvalues: $\begin{vmatrix} 2-\lambda & 2 \\ 2 & 2-\lambda \end{vmatrix} = (2-\lambda)^2 - 4 = 4 - 4\lambda + \lambda^2 - 4 = 0$

$\lambda^2 - 4\lambda = 0$

$\lambda = 4, \lambda = 0$.

$\|A\|_2 = \sigma_1 = \text{square root of largest eigenvalue of } A^TA$.

$= \sqrt{4} = 2$

6. What is the unit roundoff? (I don’t want a number, I want you to explain what the idea of the unit roundoff is.) (3 pts)

The unit roundoff is the largest relative error that can be caused by a single floating point operation (add/subtr/mult/div), and the subsequent rounding.
7. What is the singular value decomposition? Why is it useful in determining the rank of a matrix? (5 pts)

The SVD decomposes $A = U \Sigma V^T$, where $U, V$ are orthogonal matrices, and $\Sigma = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_r \end{bmatrix}$ is diagonal with $r$ non-zero entries $\sigma_i$, where $r$ is the rank of $A$.

It is useful because the calculation of the $\sigma_i$ is stable. Thus, there is usually some (numerically) that are "large," which are the real singular values, while the rest will be $\approx$ unit roundoff size, which can be ignored as numerical error. Thus, even if some are non-zero due to rounding errors, you can still more or less tell what the rank of the matrix is.

8. True/False: In descent methods, an exact line search is always better than an inexact line search. Explain why or why not. (3 pts)

False. SOR is inexact, while GS is exact, but SOR is usually better.
9. Show that orthogonal transformations preserve lengths. (First, write down what that sentence means in terms of a formula.) (5 pts)

It means \( \| Qx \|_2 = \| x \|_2 \).

\[
\text{pf: } \| Qx \|_2 = \sqrt{\langle Qx, Qx \rangle} = \sqrt{(Qx)^T Qx} = \sqrt{x^T Q^T Q x} = x^T x = \sqrt{\langle x, x \rangle} = \| x \|_2.
\]

10. Explain why Gaussian elimination with partial pivoting isn't backwards stable. Explain why we use it anyway. (5 pts)

G-E is backwards stable if \( \frac{\| L \| \| U \|}{\| A \|} \) is small. Partial pivoting guarantees that \( \| L \| \) is small, but \( \frac{\| U \|}{\| A \|} \) is still sometimes large, as large as \( 2^n \)!

We use it anyway because it almost always is backwards stable, and is computationally cheaper than complete pivoting. Plus, if we suspect a problem, we can check that \( \frac{\| L \| \| U \|}{\| A \|} \) is small for that problem.
11. Use the QR decomposition in order to solve the least squares problem

\[
\begin{bmatrix}
-4 \\
3
\end{bmatrix} [x] = [6].
\]

(You don’t have to explicitly write out the \( Q \) and \( R \), necessarily, as long as you use the method.) What is the 2-norm of the residual? (The solution/residual are simple fractions or integers.) (5 pts for solving, 1 pt for norm of residual)

\[
\begin{bmatrix}
-4 \\
3
\end{bmatrix} \Rightarrow \begin{bmatrix}
-\frac{4}{5} \\
0
\end{bmatrix}, \quad \tau = \pm 5, \quad \text{choose neg since } -4 \text{ is negative.}
\]

\[
U = \begin{bmatrix}
1/3 & -1/3
\end{bmatrix} \Rightarrow \gamma = \frac{-9}{5} = \frac{9}{5}
\]

\[
Q^T b = \begin{bmatrix}
6/3 \\
3
\end{bmatrix} - \frac{9}{5} \begin{bmatrix}
1/3 \\
-1/3
\end{bmatrix} \begin{bmatrix}
6 \\
3
\end{bmatrix}
\]

\[
\begin{bmatrix}
6/3 \\
3
\end{bmatrix} - \frac{9}{5} \begin{bmatrix}
1/3 \\
-1/3
\end{bmatrix} = \begin{bmatrix}
6 \\
3
\end{bmatrix} - \begin{bmatrix}
9/5 \\
-9/5
\end{bmatrix} = \begin{bmatrix}
-3 \\
6
\end{bmatrix}
\]

\[
\Rightarrow \begin{bmatrix}
+5 \\
0
\end{bmatrix} [x] = \begin{bmatrix}
-3 \\
6
\end{bmatrix}
\]

So \( 5x = -3 \)

\[
x = -3/5
\]

and \( \| r \|_2 = \| (6) \|_2 = 6. \)
12. Suppose \( A = \begin{bmatrix} 3 & 1 \\ 1 & 4 \end{bmatrix} \). Use one iteration of an iterative method that will approximate the eigenvalue of 
\( A \) closest to 2. Use the starting guess for an eigenvector \( q = [1; -5/8] \). About what is the eigenvalue 
of \( A \) closest to 2? (5 pts)

Use shift-and-multiply

\[
A - 2I = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \quad (A - 2I)^{-1} = \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix}.
\]

\[
(A - 2I)^{-1} \begin{bmatrix} 1 \\ -5/8 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -5/8 \end{bmatrix} = \begin{bmatrix} 2 + 5/8 \\ -1 - 5/8 \end{bmatrix} = \begin{bmatrix} 21/8 \\ -13/8 \end{bmatrix}
\]

so the eigenvalue is \( 21/8 \), for \((A - 2I)^{-1}\)

\[
8/21 \quad \text{for} \quad A - 2I
\]

\[
8/21 + 2 \quad \text{for} \quad A
\]

so \( \frac{50}{21} \) is approximately the eigenvalue of \( A \) 
closest to 2.
13. Briefly explain why the power method does not (in theory) converge to the dominant eigenvector for every possible initial vector \( q \). Then, explain why (in practice), it will anyway. (5 pts)

\[
\frac{A^j q}{\lambda_1^j} = C_1 v_1 + C_2 v_2 \frac{\lambda_2^j}{\lambda_1^j} + \cdots + C_n v_n \frac{\lambda_n^j}{\lambda_1^j}
\]

Since \( \frac{\lambda_2^j}{\lambda_1^j} \to 0 \) as \( j \to \infty \), it converges. But if \( C_1 = 0 \) to start, it can't converge to that eigenvector.

However, in practice, after a step, rounding error will cause \( C_1 \) to not quite be zero. Due to scaling so that the largest entry is 1, the other terms will still slowly diminish, and the \( v_1 \) term will be scaled up till our sequence converges.
14. For an iterative method for solving $Ax = b$ defined by a splitting $A = M - N$, with $M$ nonsingular, under what condition will this iteration converge for any choice of initial guess? Briefly explain why this condition shows the method will converge. (5 pts)

The standard iteration is $Mx^{(k+1)} = Nx^{(k)} + b$.

If $x$ is the true sol'n of $Ax = b$, then the error is $e^{(k)} = x - x^{(k)}$, and it's not too hard to see that $Me^{(k+1)} = Ne^{(k)}$.

Thus, we want $e^{(k+1)} = M^{-1}N e^{(k)} \to 0$.

Call $M^{-1}N = G$. Then,

$e^{(k)} = G^k e^{(0)} = C_1 v_1 \lambda_1^k + \ldots + C_n v_n \lambda_n^k$,

where $(v_i, \lambda_i)$ are the eigenpairs of $G$.

For this to converge, we need $|\lambda_i| < 1$ for all $i$, i.e., that $\rho(G) < 1$. 

10
Use Jacobi's.

15. Write a MATLAB program that takes as input a matrix $A$ and a vector $b$, then uses Jacobi's method or Gauss-Seidel (your choice) to solve $Ax = b$. (You can, but are not required, to have as inputs a maximum number of iterations, an initial guess and a tolerance.) You should say at the top which method your program uses. (5 pts)

```matlab
function x = Jacobi(A, b)
    n = length(b);
    x = zeros(n, 1);

    for i = 1:1000
        oldx = x;
        x = b;
        for j = 1:n
            for k = 1:j-1
                x(j) = x(j) - A(j, k) * oldx(k);
            end
            for k = j+1:n
                x(j) = x(j) - A(j, k) * oldx(k);
            end
            x(j) = x(j) / A(j, j);
        end
    end

    (you didn't have to check tolerance.)
```
16. Show that adding 3 numbers (in some order that you choose) using floating point addition is backwards stable. (5 pts)

\[
\begin{align*}
\hat{f} \left( f \left( x + y \right) + z \right) &= \left[ (x + y)(1 + \varepsilon_1) + \varepsilon_2 \right] (1 + \varepsilon_3) \\
&= x(1 + \varepsilon_1)(1 + \varepsilon_2) + y(1 + \varepsilon_1)(1 + \varepsilon_3) + z(1 + \varepsilon_2) \\
&= \hat{x} + \hat{y} + \hat{z}
\end{align*}
\]

Since \( \hat{x} = x(1 + \varepsilon_1)(1 + \varepsilon_2) \) is very close to \( x \), and similarly for the others, the approximate calculation with exact data is equivalent to an exact calculation with very close data. Thus it is backwards stable.
17. If I solve $Ax = b$ (exactly) with $A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$, and I know that my $b$ is accurate to within 0.1%, how accurate am I sure my solution $x$ is? (Use the 1 or $\infty$-norm.) (3 pts)

$$A^{-1} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}, \quad K_\infty(A) = \|A\|_\infty \|A^{-1}\|_\infty = 3 \cdot 3 = 9.$$

$$\frac{\|x\|_1}{\|x\|_1} \leq K_\infty(A) \frac{\|b\|_\infty}{\|b\|_\infty} \leq 9 \cdot 0.1\%$$

$$= 0.9\%$$

So with 0.9%.

18. Is $\|A\|_{\text{max}} = \max_{i,j} |a_{ij}|$ a matrix norm? Why or why not? (5 pts)

No. It fails submultiplicativity.

Need $\|AB\| \leq \|A\| \|B\|$.

For instance, if $A = B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, then $AB = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$.

So $\|AB\| = 2$ and $\|A\|\|B\| = 1$,

so submultiplicativity fails.
This page is also scrap paper. Make sure all final answers/important work are on the correct page.