1. Select **ALL** the answers you marked as correct. Let A be an $m \times n$ matrix. A **matrix norm** on $\mathbb{R}^{m\times n}$ is a function $\|\cdot\| : \mathbb{R}^{m\times n} \to \mathbb{R}$ satisfying the following properties:

   - ◯ $\|A\|$ is positive for all non-zero vectors, and $\|A\| = 0$ if and only if $A = 0$
   - ◯ $\|\alpha A\| = |\alpha|\|A\|$ for $A \in \mathbb{R}^{m\times n}$ and all $\alpha \in \mathbb{R}$
   - ◯ $\|A + B\| \leq \|A\| + \|B\|$ for all $A, B \in \mathbb{R}^{m\times n}$
   - ◯ $\|AB\| = \|A\|\|B\|$ for all $A, B \in \mathbb{R}^{m\times n}$

2. Let $A = \begin{pmatrix} 1 & 4 \\ 2 & 3 \\ 0 & 2 \end{pmatrix}$. Compute $\|A\|_1$, $\|A\|_\infty$, and $\|A\|_2$.

   $\|A\|_1 = max(3, 9) = 9,$
   $\|A\|_\infty = max(5, 5, 2) = 5$
   $$\begin{pmatrix} 1 & 2 & 0 \\ 4 & 3 & 2 \end{pmatrix}\begin{pmatrix} 1 & 4 \\ 2 & 3 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 5 & 10 \\ 10 & 29 \end{pmatrix} \Rightarrow \begin{pmatrix} 5-\lambda & 10 \\ 10-\lambda & 29 \end{pmatrix} \Rightarrow (5-\lambda)(29-\lambda) - 100 = 0$$
   $\Rightarrow \lambda^2 - 34\lambda + 45 = 0 \Rightarrow \lambda_1 = 17 + 2\sqrt{61}, \lambda_2 = 17 - 2\sqrt{61}$
   then, $\|A\|_2 = \sqrt{17 + 2\sqrt{61}}$

3. Let x and y be two orthogonal vectors. Then prove that $\|x + y\|_2^2 = \|x\|_2^2 + \|y\|_2^2$.

   $\|x + y\|_2^2 = (x_1 + y_1)^2 + (x_2 + y_2)^2 + ... + (x_n + y_n)^2 =$
   $(x_1^2 + y_1^2 + x_2^2 + y_2^2 + ... + x_n^2 + y_n^2) + (2x_1y_1 + 2x_2y_2 + ... + 2x_ny_n)$
   Since x, y are orthogonal, $x \cdot y = 0$, rearranging we get
   $\|x\|_2^2 + \|y\|_2^2 + 0 = \|x\|_2^2 + \|y\|_2^2$

4. Suppose x is any n vector and A is an $m \times n$ matrix. Given $\|x\|_1 \leq \sqrt{n}\|x\|_2$ and $\|x\|_2 \leq \|x\|_1$, prove that $\frac{1}{\sqrt{m}}\|A\|_1 \leq \|A\|_2$

   Since $\|A\|_1 = \max_{x \neq 0} \frac{\|Ax\|_1}{\|x\|_1}$ and $\|A\|_2 = \max_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2}$.
   Applying the vector norm inequality given, $\|Ax\|_1 \leq \sqrt{m}\|Ax\|_2$.
   Rearranging, we get $\frac{1}{\sqrt{m}}\|Ax\|_1 \leq \|Ax\|_2$ Since $\|x\|_2 \leq \|x\|_1$, we have $\frac{1}{\|x\|_1} \leq \frac{1}{\|x\|_2}$
   Then, $\frac{1}{\sqrt{m}} \frac{\|Ax\|_1}{\|x\|_1} \leq \frac{\|Ax\|_2}{\|x\|_2} \leq \max_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2} = \|A\|_2$, implying that $\frac{1}{\sqrt{m}}\|A\|_1 \leq \|A\|_2$

5. Interpret the result of of the following theorem: if $\|E\| < 1$, then $\|(I - E)^{-1} - I\| \leq \frac{\|E\|}{1 - \|E\|}$

   If the matrix E is very small, then $1 - \|E\|$ is close to unity. Thus the above result implies that if we invert a slightly perturbed identity matrix, then the error in the inverse of the perturbed matrix does not exceed the order of the perturbation.

6. Let $fl(x)$ denote the floating point representation of a real number x.
   Then, prove that the round off error is $\frac{|fl(x) - x|}{|x|} \leq \mu = \frac{1}{2}\beta^{1-t}$

   Let $x = (0.d_1 d_2 ... d_t d_{t+1} ...) \times \beta^e)$
   where $d1 \neq 0$ and $0 \leq d_i < \beta$. When we round off x we obtain one of the following floating point numbers:
   $x' = (\Delta d_1 d_2 ... d_t) \times \beta^e$
   $x'' = [(\Delta d_1 d_2 ... d_t) + \beta^{-t}] \times \beta^e$
   Assume. without any loss of generality, that x is closer to x'. We then have
   $|x - x'| \leq \frac{1}{2}|x' - x''| = \frac{1}{2}\beta^{e-t}$
   Thus, the relative error is
   $\frac{|x - x'|}{|x|} \leq (\frac{1}{2}\frac{\beta^{-t}}{\cdot d_1 d_2 ... d_t ...}) \leq \frac{1}{2}\frac{\beta^{-t}}{\frac{1}{\beta}} = \frac{1}{2}\beta^{1-t}$.

7. Fill in the blanks.

    (a) An algorithm for solving $Ax = b$ will be called **backward stable** if

    the computed solution $\hat{x}$ is an exact solution to a near-by problem $(A + E)\hat{x} = b + \delta b$ where E and δb are small.

    (b) Suppose f is a scalar-valued function with input x and $\hat{x}$ is close to x. The **absolute condition number** $C(x)$ of a function f at the point x satisfies

    $|\hat{y} - y| \approx |f'(x)||x - \hat{x}|$

    (c) The **relative condition number** $\kappa(x)$ of a function f at the point x satisfies

    $|\frac{\hat{y} - y}{y}| \approx |f'(x)\frac{x}{f(x)}|$

8. Based on your knowledge from this course, state at least two reasons why an algorithm's answer might be inaccurate.

    The problem might be ill-conditioned.
    The algorithm might be unstable.
    The test examples may be too special.
    The algorithm, though successful, might have failed in the particular circumstances.

9. Recalled that we proved the left perturbation theorem in class. Assume A is non-singular and $b \neq 0$ and $\|\Delta A\|\|A^{-1}\| < 1$, then

$$\frac{\|\delta x\|}{\|x\|} \leq (\frac{cond(A)}{1 - cond(A)\frac{\|\Delta A\|}{\|A\|}})(\frac{\|\Delta A\|}{\|A\|} + \frac{\|\delta b\|}{\|b\|})$$

    Explain that (1) what plays a crucial role in the sensitivity of the solution to perturbations in the inputs and then (2) why we can assume that $\|\Delta A\|\|A^{-1}\| < 1$.

    We see that even if the relative perturbations $\frac{\|\delta_b\|}{\|b\|}$ and $\frac{\|\Delta A\|}{\|A\|}$ small, there might he a drastic change in the solution if $Cond(A)$ is large. Thus, $Cond(A)$ plays a crucial role in the sensitivity of the solution. Because of the assumption that $\|\Delta A\| < \frac{1}{\|A^{-1}\|}$, the denominator on the right-hand side of the int:quality in this theorem is less than one. Thus even if $\frac{\|\Delta A\|}{\|A\|}$ is small, then there could be a drastic change in the solution if $Cond(A)$ is large.

10. Write a pseudocode algorithm for Gaussian Elimination without Pivoting.

   (a) Input:

   An $n \times n$ matrix (nonsingular)

   (b) Output:

   A matrix whose upper triangular part including diagonal is U and whose lower triangle is L

   (c) Pseudocode:

```
[m,n] = size(A)
if m==n
   for k = 1:(n-1)
    for i = (k+1):n
      A(i,k) = A(i,k)/A(k,k)
      A(i,k+1:n) = A(i,k+1:n) - A(i,k)*A(k,(k+1):n)
```

   (d) explain what is the purpose of pivoting.

   the purpose of pivoting is to prevent large growth in the reduced matrices, which can wipe out original data.

   (e) explain what issue might happen without pivoting, and how pivoting resolves such an issue.

   Pivoting can keep multipliers less than or equal to one in magnitude by interchange the appropriate rows/columns and make sure that the largest entry (in magnitude) is in the pivot column.

   (f) State the steps of how to use Gaussian Elimination with Partial Pivoting as matrix decomposition technique to solve for x in $Ax = b$.

   Since $PA = LU$
   $P^{-1}LUX = b$
   $LUx = Pb = b'$
   Let $Ux = y$. First solve $Ly = b'$ for y, and then solve $Ux = y$ for x.

11. Determine the following statement is true or false and give a brief explanation on your choices.

   (a) For a permutation matrix P, $P^3 = I$

   True. Because permutation matrices are orthogonal.

   (b) A permutation matrix is invertible.

   True. A permutation matrix is the same as an invertible matrix where every column is a standard basis vector.

   (c) Multiply the permutation matrix on the left interchanges columns.

   False. Multiply the permutation matrix on the left interchange rows.

   (d) The product of two permutation matrices P_1 and P_2 is a permutation matrix.

   True. Because multiply two permutation matrices only interchange rows/columns not the value of each entry.

12. Provide short answers.

   (a) Definition The growth factor ρ.

   It is the ratio of the largest element in magnitude of $A, A^{(1)}, A^{(2)}, ..., A^{(n-1)}$ to the largest element in magnitude of A.

   (b) How the growth factor could help us understand the Gaussian elimination algorithms?

   By measuring the growth of the elements in the reduced matrices $A^{(k)}$, the stability of Gaussian elimination is better understood. Also, it helps to know the round-off error property.

13. Use one iteration of iterative refinement to solve $Ax = b$ assuming $x(0) = \begin{pmatrix} 0.5 \\ 0.5 \\ 0.5 \end{pmatrix}$ where $A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix}$ and $b = \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}$.

$$r^{(0)} = b - Ax^{(0)} = \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} - \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 0.5 \\ 0.5 \\ 0.5 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} - \begin{pmatrix} 1 \\ 1.5 \\ 1.5 \end{pmatrix} = \begin{pmatrix} 0 \\ -0.5 \\ 1.5 \end{pmatrix}$$

Solve $Ac^{(0)} = r^{(0)}$

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} c_1^{(0)} \\ c_2^{(0)} \\ c_3^{(0)} \end{pmatrix} = \begin{pmatrix} 0 \\ -0.5 \\ 1.5 \end{pmatrix}$$

$$x^{(1)} = x^{(0)} + c^{(0)} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

14. State and prove one of the householder matrix's properties. (e.g. H is orthogonal. $H\mu = -\mu$. etc)

Want to Prove: H is symmetric.

Let $\beta = \dfrac{2}{\mu^T \mu}$

$H^T = (I - \beta\mu\mu^T)^T = I - \beta(\mu\mu^T)^T$

$= I - \beta(\mu^T)^T \mu^T = I - \beta\mu\mu^T = H$

Since $H^T = H$, H is symmetric.

15. Let $A = \begin{pmatrix} 1 & 4 & 2 \\ 2 & -1 & 0 \\ 1 & 1 & 3 \end{pmatrix}$, $x_0 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

(a) Choose either Rayleight Quotient or Power Method, state what you choose, and use it for 2 iterations to approximate the dominant eigenvalue of A.

Power Method:

$\hat{x}_1 = Ax_0 = \begin{pmatrix} 1 & 4 & 2 \\ 2 & -1 & 0 \\ 1 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 7 \\ 1 \\ 5 \end{pmatrix}$

$max\{|\hat{x_1}|\} = 7$, $x_1 = \begin{pmatrix} 1 \\ 1/7 \\ 5/7 \end{pmatrix}$

$\hat{x}_2 = Ax_1 = \begin{pmatrix} 1 & 4 & 2 \\ 2 & -1 & 0 \\ 1 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 1/7 \\ 5/7 \end{pmatrix} = \begin{pmatrix} 3 \\ 13/7 \\ 23/7 \end{pmatrix}$

$max\{|\hat{x_2}|\} \approx 3.3$, $x_2 \approx \begin{pmatrix} 0.9 \\ 0.56 \\ 1 \end{pmatrix}$

(b) Choose either Rayleight Quotient or Power Method which is different than your choice on part (a), write down the pseudocode.

Rayleight Quotient:
for n iterations,
compute $\omega_k = \frac{x_k^T A x_k}{x_k^T x_k}$;
solve $(A - \omega_k i)\hat{x_{k+1}}$
Normalize x_{k+1}

16. Fill in the blanks about the QR iteration.

    (a) Two matrices A and B are similar if and only if

        there exists an invertible $n \times n$ matrix X such that $xX^{-1}AX = B$

    (b) State one property of two similar matrices.

        Two similar matrices have the same eigenvalues.

    (c) During the QR iteration to find eigenvalues of A, what is the output of this algorithm?

        The output is sequence of matrices A_k containing the eigenvalues of A.

    (d) Can you make any observations on the outputs to the original matrix A?

        the A_k are similar to A.

    (e) which kind of matrix does the output converges to?

        an upper triangular matrix or to the real schur form.

17. Choose between QR Decomposition algorithm using Householder Matrices, the Householder Hessenberg Reduction algorithm, and Hessenberg QR Iteration algorithm. Write down what you choose, and the psudocode for your chosen algorithm.

    Hessenberg QR Iteration: Reduce A to upper Hessenberg via Householder Hessenberg Reduction
    Set $A_0 = A$
    for $k = 1, 2, 3...$
    find $Q_{k-1}R_{k-1}$ decomposition of A_{k-1}
    let $A_k = R_{k-1} * Q_{k-1}$

18. Briefly explain what is Least-Squares Problem.

    Given a real $m \times n$ matrix A of $rank(k) \leq min(m, n)$ and a real vector b, find a real vector x such that $r(x) = \|Ax - b\|_2$ is minimized.

19. Choose between Least Squares Solution using Normal Equations algorithm and Least-Squares Solution using Reduced QR Factorization algorithm. Write down what you choose, and the psudocode for your chosen algorithm.

    (a) Chosen:

    Least-Squares Solution using Reduced QR Factorization

    (b) Input:

    $m \times n$ matrix A; m-vector b

    (c) Output:

    least-square solution x

    (d) Psudocode:

    Find reduced QR factorization of A: $A = Q_1 R_1$
    Compute $C = Q_1^T b$
    Solve upper triangular system $R_1 x = c$.