1. Consider the following code on a machine using binary number representations:

\[
\begin{align*}
X &= 0.0 \\
10 &\ X = X + 0.1 \\
\text{PRINT } *, X, \text{ SQRT}(X) &\quad (\text{SQRT (X) is the square root of X}) \\
\text{IF (X .NE. 1.0) GO TO 10}
\end{align*}
\]

The code is trying to print out \(\sqrt{x}\) for \(x = 0.1, 0.2, \ldots, 1.0\). What problem do you expect to happen in running the code and why? Suggest a change in the code to avoid the problem.

2. Determine the linear least square approximation \(y(x) = a + bx\) to an arbitrary continuous function \(f(x)\) on \((-1, 1)\) when the inner product is defined as
\[
(f, g) = \int_{-1}^{1} f(x)g(x)dx
\]

What trouble may happen if we want to find \(a_n, a_{n-1}, \ldots, a_0\) of \(y(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0\) for large \(n\) as the least square approximation? and what is a better way to construct the least square approximation of polynomial of degree less than or equal to \(n\)?

3. Suppose we want to find solutions of the equation
\[
f(x) = x - \tan x = 0,
\]

(a) Show by using a graph that there are infinite many positive solutions to the equation.

(b) There is a root near \(3\pi/2 \approx 4.71238898\), if we take initial guess as \(x_0 = 4.7124\), and use Newton’s method, what problem do you expect to happen and why?

(c) Rearrange terms in the equation so that it is much more easier to find the solutions by Newton’s method.

4. Suppose a numerical formula \(I_h\) (like a numerical integration formula) with step size \(h\) is used to approximate a mathematical expression \(I\) (like a definite integral). If the error of the formula is given by
\[
I_h - I = kh^p + O(h^{p+2}), \quad \text{where } k, p \text{ are constants}
\]

(a) describe Richardson extrapolation which uses \(I_h, I_{h/2}\) to generate a more accurate numerical formula \(\tilde{I}_{h/2}\).

(b) Apply Richardson extrapolation to the trapezoidal rule
\[
I(f) = \int_{a}^{b} f(x)dx \approx I_h(f) = \frac{h}{2}(f(a) + f(b)), \quad h = b - a
\]

to derive a more accurate integration formula. Identify this more accurate integration formula (find the familiar name of the formula). (Hint: \(I_{h/2}\) would use two subintervals)

5. Let \(A\) be a real \(n \times n\) matrix whose eigenvalues satisfy \(0 < \lambda_n < \lambda_{n-1} < \cdots < \lambda_1\). State and prove convergence of a numerical method for finding \(\lambda_1\) and \(\lambda_n\).
6. Show that if \( A \in \mathbb{R}^{m \times n} \) has rank \( n \), the \( \| A(A^T A)^{-1} A^T \|_2 = 1 \), where \( A^T \) is the transpose of \( A \).

7. Suppose \( A \in \mathbb{R}^{n \times n} \), \( A^T \) (the transpose of \( A \)) is diagonally dominant, i.e.,
\[
|a_{ii}| \geq \sum_{i \neq j} |a_{ij}|
\]
and \( A \) is nonsingular, show that \( A = LU \) with \( L \) being a unit lower triangular matrix, i.e., Gauss elimination can be performed without pivoting, and \( |l_{ij}| \leq 1 \), where \( l_{ij} \) are entries in \( L \).
(Hint: consider a partition of \( A \) of the form:
\[
A = \begin{bmatrix} \alpha & w^T \\ v & B \end{bmatrix}, \text{where } B \in \mathbb{R}^{(n-1) \times (n-1)}, v, w \in \mathbb{R}^{n-1}.
\]
and consider one step of Gauss elimination)

8. Given \( A \in \mathbb{R}^{n \times n} \), a symmetric positive matrix, solving the linear system \( Ax = b \) for \( x \in \mathbb{R}^n \) is equivalent to minimizing the functional
\[
\phi(x) = \frac{1}{2} x^T A x - x^T b, \quad \text{where } x^T \text{ is the transpose of } x
\]
Suppose
\[
P_k = [p_1, p_2, \ldots, p_k] \in \mathbb{R}^{n \times k}, p_i \in \mathbb{R}^n, i = 1, 2, \ldots, k
\]
if \( x \in \text{span} \{p_1, p_2, \ldots, p_k\} \), then
\[
x = P_{k-1} y + \alpha p_k, \quad P_{k-1} = [p_1, \ldots, p_{k-1}], y \in \mathbb{R}^{k-1}, \alpha \in \mathbb{R}.
\]
It can be derived that
\[
\phi(x) = \frac{1}{2} \phi(P_{k-1} y) + \alpha y^T P_{k-1}^T A p_k + \frac{\alpha^2}{2} p_k^T A p_k - \alpha p_k^T b
\]
The Conjugate Gradient method can be constructed as follows:
\[
k = 0; x_0 = 0; r_0 = b, \text{ }(r = b - Ax \text{ is the residual}),
\]
while
\[
r_k \neq 0
\]
\[
k = k + 1
\]
if \( k = 1, p_1 = r_0 \)
otherwise choose \( p_k \in \text{span}\{Ap_1, Ap_2, \ldots, Ap_{k-1}\}^\perp \), such that \( p_k^T r_{k-1} \neq 0 \)
\[
\alpha_k = p_k^T r_{k-1} / p_k^T A p_k
\]
\[
x_k = x_{k-1} + \alpha_k p_k
\]
\[
r_k = b - A x_k
\]
end
Show that in the algorithm, \( x_k \) minimizes the functional \( \phi(x) \) over \( \text{span} \{p_1, p_2, \ldots, p_k\} \). Furthermore \( p_i^T r_k = 0, i = 1, 2, \ldots, k \).