Introduction

Numerical analysis is concerned with the design and analysis of algorithms for solving mathematical problems that arise in many fields, especially science and engineering. Scientific computing could be regarded as a combination of modeling, visualization, and numerical analysis. In particular, it deals with quantities that are continuous, as opposed to discrete which occurred in most other parts of computer science.

- Sources of Approximation
  - Modeling
  - Empirical measurements
  - Previous computations
  - Discretization
  - Truncation and Rounding

- Types of Errors
  - Absolute Error and Relative Error
  - Data Error and Computational Error
  - Truncation Error and Rounding Error
  - Sensitivity and Condition Number
  - Stability and Accuracy

- Computer Arithmetic
  - Floating-Point Numbers
  - Representation and Normalization
  - Machine Precision
  - Summing a series $\sum_{n=1}^{\infty} \frac{1}{n}$
  - Quadratic formula (Solving $ax^2 + bx + c = 0$)
• Mathematical Software and Environments
  Reliability, Accuracy, Portability
  Robustness, Efficiency, Maintainability
  Usability, Applicability
  Matlab, IMSL (PV-Wave), Netlib, NAG

• Historical Notes from Mathematicians
  Newton (1642~1727), Euler (1707~1783), Lagrange (1736~1813)
  Laplace (1749~1827), Legendre (1752~1833), Gauss (1777~1855)
  Cauchy (1789~1857), Jacobi (1804~1851), Adams (1819~1892),
  Chebyshev (1821~1894), Hermite (1822~1901), Laguerre (1834~1886)
Types of Errors

Let $y$ be the true value, and let $\hat{y}$ be the approximate value, then

Absolute error $= \hat{y} - y$

Relative error $= \frac{\hat{y} - y}{y}$

Let $f : R \rightarrow R$. Suppose that we must work with inexact input $\hat{x}$ and we can compute only an approximation to the function $\hat{f}$. Then, we have

Total error $= \hat{f}(\hat{x}) - f(x) = [\hat{f}(\hat{x}) - f(\hat{x})] + [f(\hat{x}) - f(x)]$

Computation error $= \hat{f}(\hat{x}) - f(\hat{x})$

Propagated data error $= f(\hat{x}) - f(x)$

Computing $\sin(\pi/8)$

Use $\hat{f}(x) = x$ instead of $f(x) = \sin(x)$ and use $\hat{x} = 3/8$ to approximate $x = \pi/8$. By calculator, we have

$f(\pi/8) = \sin(\pi/8) \approx 0.3827$ and $f(3/8) \approx 0.3663$

On the other hand,

$\hat{f}(\hat{x}) - f(\hat{x}) = \hat{x} - \sin(3/8) = 0.3750 - 0.3663 = 0.0087$

$f(\hat{x}) - f(x) = 0.3663 - 0.3827 = -0.0164$

So the total error is

$\hat{f}(\hat{x}) - f(x) = [\hat{f}(\hat{x}) - f(\hat{x})] + [f(\hat{x}) - f(x)] = 0.0087 + (-0.0164) = -0.0077$

$\sqrt{3.14159} \approx 3.1415$ by truncation, $3.14159 \approx 3.1416$ by round-off.

$f(x + h) = f(x) + f'(x)h + f''(\theta)h^2/2$ for some $\theta \in [x, x + h]$
Suppose that we want to compute \( y = f(x) \), where \( f : R \to R \), but we obtain instead an approximate value \( \hat{y} \). The discrepancy \( \Delta y = \hat{y} - y \) is called the forward error. On the other hand, how much data error in the initial input would be required to explain all of the error in the final computed result? The quantity \( \Delta x = \hat{x} - x \), where \( f(\hat{x}) = \hat{y} \), is called backward error.

Consider \( f(x) = \sqrt{x} \), as an approximation to \( y = \sqrt{2} \), \( \hat{y} = 1.4 \) has an absolute forward error \( |\Delta y| = |1.4 - 1.4142 \cdots | \). The backward error \( |\Delta x| = |\hat{x} - x| = |(1.4)^2 - 2| = 0.04 \). So, \( |\Delta y|/y \approx 1\% \) and \( |\Delta x|/x \approx 2\% \).

A problem is said to be insensitive, or well-conditioned, if a given relative change in the input data causes a reasonably commensurate relative change in the solution. Otherwise, it is called sensitive, or ill-conditioned.

\[
\text{Condition number} = \frac{|\Delta y/y|}{|\Delta x/x|} = \frac{|[f(\hat{x}) - f(x)]/f(x)|}{|\hat{x} - x|/x}
\]

If \( f \) is continuously differentiable around \( x \), then

\[
\text{Condition number} = \left| \frac{xf'(x)}{f(x)} \right|
\]

⊗ Implement and discuss the following problems.

(1) \( \sum_{n=1}^{\infty} \frac{1}{n} \) (Prob. 1.8 on p.45)

(2) Solve \( x^3 + ax^2 + bx + c = 0 \) (Prob. 1.11 on p.46)

(3) Evaluate \( \|x\|_2 = (\sum_{i=1}^{n} x_i^2)^{1/2} \) (Prob. 1.15 on p.47)
Norms of Vectors and Matrices

Definition: A vector norm on $R^n$ is a function

$$\tau : R^n \rightarrow R^+ = \{x \geq 0 \mid x \in R\}$$

that satisfies

1. $\tau(x) > 0 \ \forall \ x \neq 0, \ \tau(0) = 0$
2. $\tau(cx) = |c|\tau(x) \ \forall \ c \in R, \ x \in R^n$
3. $\tau(x + y) \leq \tau(x) + \tau(y) \ \forall \ x, y \in R^n$

Hölder norm ($p$-norm) $\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}$ for $p \geq 1$.

- ($p=1$) $\|x\|_1 = \sum_{i=1}^{n} |x_i|$ (Mahattan or City-block distance)
- ($p=2$) $\|x\|_2 = \left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2}$ (Euclidean distance)
- ($p=\infty$) $\|x\|_\infty = \max_{1 \leq i \leq n} \{|x_i|\}$ ($\infty$-norm)
Definition: A matrix norm on $\mathbb{R}^{m \times n}$ is a function
\[
\tau : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^+ = \{ x \geq 0 \mid x \in \mathbb{R} \}
\]
that satisfies

1. $\tau(A) > 0 \quad \forall \ A \neq O$, $\tau(O) = 0$
2. $\tau(cA) = |c|\tau(A) \quad \forall \ c \in \mathbb{R}, \ A \in \mathbb{R}^{m \times n}$
3. $\tau(A + B) \leq \tau(A) + \tau(B) \quad \forall \ A, B \in \mathbb{R}^{m \times n}$

Consistency Property: $\tau(AB) \leq \tau(A)\tau(B) \quad \forall \ A, B$

(a) $\tau(A) = max\{|a_{ij}| \mid 1 \leq i \leq m, \ 1 \leq j \leq n\}$
(b) $\|A\|_F = \left[\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2\right]^{1/2}$ (Fröbenius norm)

Subordinate Matrix Norm: $\|A\| = max_{\|x\| \neq 0}\{\|A\|/\|x\|\}$

1. If $A \in \mathbb{R}^{m \times n}$, then $\|A\|_1 = max_{1 \leq j \leq n} (\sum_{i=1}^{m} |a_{ij}|)$
2. If $A \in \mathbb{R}^{m \times n}$, then $\|A\|_\infty = max_{1 \leq i \leq m} (\sum_{j=1}^{n} |a_{ij}|)$
3. Let $A \in \mathbb{R}^{n \times n}$ be real symmetric, then $\|A\|_2 = max_{1 \leq i \leq n} |\lambda_i|$, where $\lambda_i \in \lambda(A)$
Problems Solved by Matlab

Let \( A, B, H, x, y, u, b \) be matrices and vectors defined below, and \( H = I - 2uu^t \)

\[
A = \begin{bmatrix}
2 & 1 & 1 \\
4 & 6 & 0 \\
-2 & 7 & 2
\end{bmatrix}, \quad
B = \begin{bmatrix}
-3 & 1 & 0 \\
1 & -3 & 0 \\
0 & 0 & 3
\end{bmatrix}, \quad
u = \begin{bmatrix}
1/2 \\
1/2 \\
1/2
\end{bmatrix}, \quad
b = \begin{bmatrix}
6 \\
2 \\
-5
\end{bmatrix}, \quad
x = \begin{bmatrix}
1 \\
1 \\
1
\end{bmatrix}, \quad
y = \begin{bmatrix}
1 \\
1 \\
-1
\end{bmatrix}
\]

1. Let \( A=LU=QR \), find \( L, U, Q, R \).
2. Find determinants and inverses of matrices \( A, B, \) and \( H \).
3. Solve \( Ax = b \), how to find the number of floating-point operations are required?
4. Find the ranks of matrices \( A, B, \) and \( H \).
5. Find the characteristic polynomials of matrices \( A \) and \( B \).
6. Find 1-norm, 2-norm, and \( \infty \)-norm of matrices \( A, B, \) and \( H \).
7. Find the eigenvalues/eigenvectors of matrices \( A \) and \( B \).
8. Find matrices \( U \) and \( V \) such that \( U^{-1}AU \) and \( V^{-1}BV \) are diagonal matrices.
9. Find the singular values and singular vectors of matrices \( A \) and \( B \).
10. Randomly generate a \( 4 \times 4 \) matrix \( C \) with \( 0 \leq C(i,j) \leq 9 \).