A FIRST COURSE

IN

NUMERICAL ANALYSIS

A FIRST COURSE

IN

NUMERICAL ANALYSIS

MAT4001 Notebook

Prof. Yutian Li

The Chinese University of Hongkong, Shenzhen

香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen

Contents

Acknowledgments

This book is from the MAT4001 in fall semester, 2018.

CUHK(SZ)

Notations and Conventions

$\mathbb{R}^n$	n-dimensional real space
$\mathbb{C}^n$	n-dimensional complex space
$\mathbb{R}^{m\times n}$	set of all $m \times n$ real-valued matrices
$\mathbb{C}^{m\times n}$	set of all $m \times n$ complex-valued matrices
x_i	ith entry of column vector $\boldsymbol{x}$
a_{ij}	(i,j)th entry of matrix $\boldsymbol{A}$
$\boldsymbol{a}_i$	ith column of matrix $\boldsymbol{A}$
$\boldsymbol{a}_i^{\mathrm{T}}$	ith row of matrix $\boldsymbol{A}$
$\mathbb{S}^n$	set of all $n \times n$ real symmetric matrices, i.e., $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ and $a_{ij} = a_{ji}$ for all i,j
$\mathbb{H}^n$	set of all $n \times n$ complex Hermitian matrices, i.e., $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ and $\bar{a}_{ij} = a_{ji}$ for all i,j
$\boldsymbol{A}^{\mathrm{T}}$	transpose of $\boldsymbol{A}$, i.e, $\boldsymbol{B} = \boldsymbol{A}^{\mathrm{T}}$ means $b_{ji} = a_{ij}$ for all i,j
$\boldsymbol{A}^{\mathrm{H}}$	Hermitian transpose of $\boldsymbol{A}$, i.e, $\boldsymbol{B} = \boldsymbol{A}^{\mathrm{H}}$ means $b_{ji} = \bar{a}_{ij}$ for all i,j
$\mathrm{trace}(\boldsymbol{A})$	sum of diagonal entries of square matrix $\boldsymbol{A}$
$\mathbf{1}$	A vector with all 1 entries
$\mathbf{0}$	either a vector of all zeros, or a matrix of all zeros
$\boldsymbol{e}_i$	a unit vector with the nonzero element at the ith entry
$\mathcal{C}(\boldsymbol{A})$	the column space of $\boldsymbol{A}$
$\mathcal{R}(\boldsymbol{A})$	the row space of $\boldsymbol{A}$
$\mathcal{N}(\boldsymbol{A})$	the null space of $\boldsymbol{A}$
$\mathrm{Proj}_{\mathcal{M}}(\boldsymbol{A})$	the projection of $\boldsymbol{A}$ onto the set $\mathcal{M}$

Chapter 1

Week1

1.1. Monday

1.1.1. Introduction to Numerical Analysis

Solving Nonlinear Equations. For example, we want to solve a nonlinear equation $w(1)$ with w to be the LambertW function:

$$we^{w} = 1.$$

This topic will be taught in chapter 2.

Interpolation. Given a list of data points, our aim is to recover/approximate the origin function over a function class, i.e., piecewise linear functions or polynomials. This topic will be taught in chapter 3.

Numerical Integration. The cdf of the standard normal distribution is given by:

$$N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{t^2}{2}\right) \mathrm{d}t$$

To approximate the values for $N(x)$, we need to apply a numerical method (quadrature rules) to evaluate. This topic will be taught in chapter 4.

Solving Linear Systems. To find the solutions of a linear system of equations

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},$$

e.g., when we use *finite difference* method to solve a differential equation, it is necessary to apply some numerical method to solve it in computer. This topic will be taught in chapter 5.

Least Squares. If we have more time, we will teach how to fit a set of data points by a function from a function class.

1.1.2. Basic Concepts

> **Definition 1.1** [Truncation Error] The error made by numerical algorithms that arises from taking finite number of steps in computation ■

For example, consider the taylor's theorem

$$f(x) = P_n(x) + R_n(x)$$

where

$$\begin{aligned} P_n(x) &= f(x_0) + f'(x_0)(x - x_0) + \cdots + \frac{f^{(n))}(x_0)}{n!}(x - x_0)^n \\ R_n(x) &= \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)^{n+1} \end{aligned}$$

If we use $P_n(x)$ to approximate $f(x)$, the error $R_n(x)$ is called the **truncation error**.

> **Definition 1.2** [Round-off Error] The error produced when a computer is used to perform real number calculations, i.e., the computer only gives approximate value for some real numebers. ■

For example, for numbers $\frac{1}{3}, \pi, \sqrt{2}$, they cannot be represented exactly in the computation by a computer.

Such errors are invertible, but we can try to minimize the negative impact of these errors by rewriting the formula that are wish to compute.

Definition 1.3 [Binary Floating Point Number] A 64-bit (binary digit) representation is used for a real number:

$$(-1)^s 2^{c-1023}(1+f)$$

■

Definition 1.4 Suppose p^* is an approximation to p. The acutal error is $p - p^*$, the absolute error is $|p - p^*|$, and the relative error is $\frac{|p-p^*|}{|p|}$ provided that $p \neq 0$. ■

Definition 1.5 Suppose p^* is an approximation to p, then p^* is said to approximate p to t **significant digits** if t is the largest nonnegative integer for which

$$\frac{|p - p^*|}{|p|} \leq 5 \times 10^{-t}.$$

■

1.1.3. Convergence and Stability

Definition 1.6 [Convergence] Suppose a sequence $\{\beta_n\}_{n=1}^{\infty}$ is known to converge to zero, and $\{\alpha_n\}_{n=1}^{\infty}$ converges to a number α. If there exists $K > 0$ such that

$$|\alpha_n - \alpha| \leq K|\beta_n| \quad \text{for large } n,$$

then $\{\alpha_n\}_{n=1}^{\infty}$ is said to converge to α with rate of convergence $\mathcal{O}(\beta_n)$, which is denoted as

$$\alpha_n = \alpha + \mathcal{O}(\beta_n).$$

■