Chapter 1
Mathematical Preliminaries and Error Analysis

Per-Olof Persson
persson@berkeley.edu

Department of Mathematics
University of California, Berkeley

Math 128A Numerical Analysis

Limits and Continuity

Definition

A function f defined on a set X of real numbers has the *limit* L at x_0, written $\lim_{x \to x_0} f(x) = L$, if, given any real number $\varepsilon > 0$, there exists a real number $\delta > 0$ such that

$$|f(x) - L| < \varepsilon, \quad \text{whenever} \quad x \in X \text{ and } 0 < |x - x_0| < \delta.$$

Definition

Let f be a function defined on a set X of real numbers and $x_0 \in X$. Then f is *continuous* at x_0 if

$$\lim_{x \to x_0} f(x) = f(x_0).$$

The function f is continuous on the set X if it is continuous at each number in X.

Limits of Sequences

Definition

Let $\{x_n\}_{n=1}^{\infty}$ be an infinite sequence of real of complex numbers. The sequence $\{x_n\}_{n=1}^{\infty}$ has the *limit* x if, for any $\varepsilon > 0$, there exists a positive integer $N(\varepsilon)$ such that $|x_n - x| < \varepsilon$, whenever $n > N(\varepsilon)$. The notation

$$\lim_{n\to\infty} x_n = x, \text{ or } x_n \to x \text{ as } n \to \infty,$$

means that the sequence $\{x_n\}_{n=1}^{\infty}$ converges to x.

Theorem

If f is a function defined on a set X of real numbers and $x_0 \in X$, then the following statements are equivalent:

1. f is continuous at x_0;
2. If the sequence $\{x_n\}_{n=1}^{\infty}$ in X converges to x_0, then $\lim_{n\to\infty} f(x_n) = f(x_0)$.

Derivatives

Definition

Let f be a function defined in an open interval containing x_0. The function f is *differentiable* at x_0 if

$$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

exists. The number $f'(x_0)$ is called the *derivative* of f at x_0. A function that has a derivative at each number in a set X is *differentiable* on X.

Theorem

If the function f is differentiable at x_0, then f is continuous at x_0.

Derivative Theorems

Theorem (Rolle's Theorem)

Suppose $f \in C[a, b]$ and f is differentiable on (a, b). If $f(a) = f(b)$, then a number c in (a, b) exists with $f'(c) = 0$.

Theorem (Mean Value Theorem)

If $f \in C[a, b]$ and f is differentiable on (a, b), then a number c in (a, b) exists with

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Theorem (Extreme Value Theorem)

If $f \in C[a, b]$, then $c_1, c_2 \in [a, b]$ exist with $f(c_1) \leq f(x) \leq f(c_2)$, for all $x \in [a, b]$. In addition, if f is differentiable on (a, b), then the numbers c_1 and c_2 occur either at the endpoints of $[a, b]$ or where f' is zero.

Integrals

Definition

The *Riemann integral* of the function f on the interval $[a, b]$ is the following limit, provided it exists:

$$\int_a^b f(x)\,dx = \lim_{\max \Delta x_i \to 0} \sum_{i=1}^{n} f(z_i)\Delta x_i,$$

where the numbers $x_0, x_1, \ldots, x_n$ satisfy $a = x_0 \leq x_1 \leq \cdots \leq x_n = b$, and where $\Delta x_i = x_i - x_{i-1}$, for each $i = 1, 2, \ldots, n$, and z_i is arbitrarily chosen in the interval $[x_{i-1}, x_i]$.

Theorem (Weighted Mean Value Theorem for Integrals)

Suppose $f \in C[a, b]$, the Riemann integral of g exists on $[a, b]$, and $g(x)$ does not change sign on $[a, b]$. Then there exists a number c in (a, b) with

$$\int_a^b f(x)g(x)\,dx = f(c)\int_a^b g(x)\,dx.$$

Generalizations

Theorem (Generalized Rolle's Theorem)

Suppose $f \in C[a, b]$ is n times differentiable on (a, b). If $f(x)$ is zero at the $n + 1$ distinct numbers $x_0, \ldots, x_n$ in $[a, b]$, then a number c in (a, b) exists with $f^{(n)}(c) = 0$.

Theorem (Intermediate Value Theorem)

If $f \in C[a, b]$ and K is any number between $f(a)$ and $f(b)$, then there exists a number c in (a, b) for which $f(c) = K$.

Taylor Polynomials

Theorem (Taylor's Theorem)

Suppose $f \in C^n[a,b]$, that $f^{(n+1)}$ exists on $[a,b]$, and $x_0 \in [a,b]$. For every $x \in [a,b]$, there exists a number $\xi(x)$ between x_0 and x with $f(x) = P_n(x) + R_n(x)$, where

$$P_n(x) = f(x_0) + f'(x_0)(x-x_0) + \frac{f''(x_0)}{2!}(x-x_0)^2 + \\ + \cdots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k$$

and

$$R_n(x) = \frac{f^{(n+1)}(\xi(x))}{(n+1)!}(x-x_0)^{n+1}.$$

IEEE Floating Point Numbers

Long real (double precision) format

- Widely adopted standard
- Default data type in MATLAB, “`double`” in C
- Base 2, 1 sign bit, 11 exponent bits, 52 significand bits:

x	xxxxxxxxxxx	xxx $\cdots$ (52 total bits) $\cdots$ xxx
s	c	f

- Represented number:

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

Decimal Floating-Point Numbers

Base-10 Floating-Point

- For simplicity we study k-digit *decimal machine numbers*:

$$\pm 0.d_1 d_2 \dots d_k \times 10^n, \quad 1 \le d_1 \le 9, \quad 0 \le d_i \le 9$$

- Any positive number within the range can be written:

$$y = 0.d_1 d_2 \cdots d_k d_{k+1} d_{k+2} \dots \times 10^n$$

- Two ways to represent y with k digits:
  - *Chopping*: Chop off after k digits:

$$fl(y) = 0.d_1 d_2 \dots d_k \times 10^n$$

  - *Rounding*: Add $5 \times 10^{n-(k+1)}$ and chop:

$$fl(y) = 0.\delta_1 \delta_2 \dots \delta_k \times 10^n$$

Errors and Significant Digits

Definition

If p^* is an approximation to p, the *absolute error* is $|p - p^*|$, and the *relative error* is $|p - p^*|/|p|$, provided that $p \neq 0$.

Definition

The number p^* is said to approximate p to t *significant digits* (or figures) if t is the largest nonnegative integer for which

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

Floating Point Operations

Finite-Digit Arithmetic

- Machine addition, subtraction, multiplication, and division:

$$x \oplus y = fl(fl(x) + fl(y)), \quad x \otimes y = fl(fl(x) \times fl(y))$$
$$x \ominus y = fl(fl(x) - fl(y)), \quad x \oslash y = fl(fl(x) / fl(y))$$

- "Round input, perform exact arithmetic, round the result"

Cancellation

- Common problem: Subtraction of nearly equal numbers:

$$fl(x) = 0.d_1 d_2 \dots d_p \alpha_{p+1} \alpha_{p+2} \dots \alpha_k \times 10^n$$
$$fl(y) = 0.d_1 d_2 \dots d_p \beta_{p+1} \beta_{p+2} \dots \beta_k \times 10^n$$

gives fewer digits of significance:

$$fl(fl(x) - fl(y)) = 0.\sigma_{p+1} \sigma_{p+2} \dots \sigma_k \times 10^{n-p}$$

Error Growth and Stability

Definition

Suppose $E_0 > 0$ is an initial error, and E_n is the error after n operations.

- $E_n \approx CnE_0$: *linear* growth of error
- $E_n \approx C^n E_0$: *exponential* growth of error

Stability

- *Stable* algorithm: Small changes in the initial data produce small changes in the final result
- *Unstable* or *conditionally stable* algorithm: Large errors in final result for all or some initial data with small errors

Rate of Convergence (Sequences)

Definition

Suppose $\{\beta_n\}_{n=1}^{\infty}$ is a sequence converging to zero, and $\{\alpha_n\}_{n=1}^{\infty}$ converges to a number α. If a positive constant K exists with

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

then we say that $\{\alpha_n\}_{n=1}^{\infty}$ converges to α with *rate of convergence* $O(\beta_n)$, indicated by $\alpha_n = \alpha + O(\beta_n)$.

Polynomial rate of convergence

- Normally we will use

$$\beta_n = \frac{1}{n^p},$$

  and look for the largest value $p > 0$ such that $\alpha_n = \alpha + O(1/n^p)$.

Rate of Convergence (Functions)

Definition

Suppose that $\lim_{h\to 0} G(h) = 0$ and $\lim_{h\to 0} F(h) = L$. If a positive constant K exists with

$$|F(h) - L| \le K|G(h)|, \quad \text{for sufficiently small } h,$$

then we write $F(h) = L + O(G(h))$.

Polynomial rate of convergence

- Normally we will use

$$G(h) = h^p,$$

  and look for the largest value $p > 0$ such that $F(h) = L + O(h^p)$.