Computing Method

Recitation Notes

Yu Xinrui

November 5, 2022

Knowledge Overview

- 1 Basic Notes about Arithmetic
- 2 Approximating Functions
  - Interpolation
    - Lagrange Interpolation
    - Newton Interpolation
    - Hermite Interpolation
    - spline
  - Fitting (Least-square fitting)
- 3 Solution of non-linear equation
  - Bisection
  - Newton's Iteration
  - Secant Method
  - Fixed-Point Iteration

Knowledge Overview

- 4 Solution of linear systems of equations
  - Elimination and Decomposition
    - Gauss Elimination
    - LU Decomposition
  - Iteration
    - Richardson
    - Jacobi
    - Gauss-Seidel
    - SOR
- 5 Numerical Differentiation and Integration
  - Numerical Differentiation
  - Newton Cotes
  - Gauss Integration

- Absolute Error and Relative Error

Definition

Let x be a real number and x^* is its approximation, then the error is $x - x^*$. The absolute error is $|x - x^*|$ The relative error is $|\frac{x-x^*}{x}|$

- Significant Digits

Definition

If the absolute error of x is no more than half of a certain digit, then the total number of digits from this digit to the first non-zero digit is called significance digits.

e.g.

1.Basic Notes about Arithmetic

Principle of value compututation:

1. Avoid subtraction of nearly equal quantities.
2. Minimize number of calculation.
3. Avoid extremely small denominator.

Frequently used methods:

1. Taylor expansion
2. Sum to product formula
3. Qin Jiu Shao Algorithm:

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

$$=(\cdots(a_n x + a_{n-1})x + \cdots + a_1)x + a_0$$

2.Approximating Functions-Interpolation

Theorem on Polynomial Intepolation

If $x_0, \cdots, x_n$ are distinct real numbers, then for arbitrary values $y_0, \cdots, y_n$, there is a unique polynomial p_n of degree at most n such that $p_n(x) = y_i, 0 \le i \le n$

- Lagrange intepolation

A table of data pairs $(x_i, y_i), 0 \le i \le n$, the Lagrange interpolation polynomial of order n has the form:

$$p(x) = y_0 l_0(x) + y_1 l_1(x) + \cdots + y_n l_n(x) = \sum_{i=0}^{n} y_i l_i(x)$$

where the lagrange bases function :

$$l_i(x) = \prod_{j=0, j\neq i}^{n} \frac{x - x_j}{x_i - x_j}$$

- Newton intepolation

$$p(x) = \sum_{k=0}^{n} f[x_0, x_1, \cdots, x_k] \prod_{j=0}^{k-1} (x - x_j)$$

- To calculate Newton interpolation polynomial, use Divided Difference Method

Theorem on Polynomial Intepolation Error

Let f be a function in $C^{n+1}[a,b]$, and let p be the polynomial of degree at most n that intepolates the function f at n+1 distinct points $x_0, \cdots, x_n$ in the interval [a,b]. To each x in [a,b] there corresponds a point ξ_x in (a,b) such that

$$f(x) - p(x) = \frac{1}{(n+1)!} f^{(n+1)}(\xi_x) \prod_{i=0}^{n} (x - x_i) \tag{1}$$

- Hermite interpolation

Theorem on Hermite Interpolation

Let the nodes be $x_0, x_1, \cdots, x_n$, and suppose that at node x_i these interpolation condition are given:

$$p^{(j)}(x_j) = c_{ij}, (0 \leq j \leq k_i - 1, = \leq j \leq n)$$

The total number of conditions on p is denoted by m+1, and therefore

$$m + 1 = k_0 + k_1 + \cdots, k_n$$

Then there exists a unique polynomial p in $\prod_m$ satisfying the Hermite interpolation conditions.

- To calculate Hermite interpolation polynomial, use Divided Difference Method

- 3-spline

Suppose there are n+1 nodes on [a,b], to derive cubic polynomial on n intervals, we need 4n conditions. By definition of spline interpolation, we still need n+1 function values and 2 boundary conditions. To calculate M_i on n-1 inner nodes, there are two common cases:

(1) 给定 M_0, M_n 的值，此时 $n-1$ 个方程组有 $n-1$ 个未知量 $\{M_i, i=1, 2, \cdots, n-1\}$. 当 $M_0=0, M_n=0$ 时，称为自然边界条件.

$$\begin{bmatrix} 2 & \lambda_1 & & & \\ \mu_2 & 2 & \lambda_2 & & \\ & \ddots & \ddots & \ddots & \\ & & \lambda_{n-2} & 2 & \lambda_{n-2} \\ & & & \mu_{n-1} & 2 \end{bmatrix} \begin{bmatrix} M_1 \\ M_2 \\ \vdots \\ M_{n-2} \\ M_{n-1} \end{bmatrix} = \begin{bmatrix} d_1 - \mu_1 M_0 \\ d_2 \\ \vdots \\ d_{n-2} \\ d_{n-1} - \lambda_{n-1} M_n \end{bmatrix}$$

(2) 给定 $S'(x_0)=m_0, S'(x_n)=m_n$ 的值，将 $S'(x_0)=m_0, S'(x_n)=m_n$ 的值分别代入 $S'(x)$ 在 $[x_0, x_1], [x_{n-1}, x_n]$ 中的表达式，得到另外两个方程:

$$2M_0 + M_1 = \frac{6}{h_0}[f[x_0, x_1] - m_0] = d_0$$

$$M_{n-1} + 2M_n = \frac{6}{h_{n-1}}[m_n - f[x_{n-1}, x_n]] = d_n$$

得到 $n+1$ 个未知量, $n+1$ 个方程组

$$\begin{bmatrix} 2 & 1 & & & & \\ u_1 & 2 & \lambda_1 & & & \\ & u_2 & 2 & \lambda_2 & & \\ & & \ddots & \ddots & \ddots & \\ & & & u_{n-2} & 2 & \lambda_{n-1} \\ & & & & 1 & 2 \end{bmatrix} \begin{bmatrix} M_0 \\ M_1 \\ M_2 \\ \vdots \\ M_{n-1} \\ M_n \end{bmatrix} = \begin{bmatrix} d_0 \\ d_1 \\ d_2 \\ \vdots \\ d_{n-1} \\ d_n \end{bmatrix}$$

$$\lambda_i = \frac{h_i}{h_i + h_{i-1}}, \quad \mu_i = 1 - \lambda_i$$

$$d_i = \frac{6}{h_i + h_{i-1}} \left(\frac{y_{i+1} - y_i}{h_i} - \frac{y_i - y_{i-1}}{h_{i-1}} \right) = 6f(x_{i-1}, x_i, x_{i+1})$$

Then, use

$$S_i(x) = \frac{(x_{i+1} - x)^3}{6h_i} M_i + \frac{(x - x_i)^3}{6h_i} M_{i+1} + c(x_{i+1} - x) + d(x - x_i)$$

where c=$\frac{y_i}{h_i} - \frac{h_i M_i}{6}$, d=$\frac{y_{i+1}}{h_i} - \frac{h_i M_{i+1}}{6}$,$i = 0, 1, \cdots, n-1$ to obtain interpolation polynomials.

- Least-Square problem on function fitting

Definition

Let f(x) is a function on [a,b], $x_{i}{}_{i=0}^{m}$ be m+1 distinct nodes on the [a,b],Φ be a given class of functions. Find function ϕ on Φ such that f(x) and $\phi(x)$ are closest on the m+1 nodes. If the distance is measured by 2-norm, then it is least-square problem. i.e. $R = \sqrt{\sum_{i=0}^{m}(\phi(x_i) - f(x_i))^2}$ is minimized.

- If Φ is polynomial space, it is the well-known polynomial fitting.

To find fitting polynomial in this way, we form the normal equation:

$$\begin{pmatrix} m & \sum_{i=1}^{m} x_i & \cdots & \sum_{i=1}^{m} x_i^n \\ \sum_{i=1}^{m} x_i & \sum_{i=1}^{m} x_i^2 & \cdots & \sum_{i=1}^{m} x_i^{n+1} \\ \vdots & \vdots & & \vdots \\ \sum_{i=1}^{m} x_i^n & \sum_{i=1}^{m} x_i^{n+1} & \cdots & \sum_{i=1}^{m} x_i^{2n} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{m} y_i \\ \sum_{i=1}^{m} x_i y_i \\ \vdots \\ \sum_{i=1}^{m} x_i^n y_i \end{pmatrix}$$

2.Approximating Functions-Fitting

- Contradictory Equations and least-squqre solution

Explanation

A system of m equations of n unknows is:$A\alpha = Y$, if $m \geq n$, then it has no solution. The least-square solution of it is the x that minimizes $\|Y - A\alpha\|^2$. Such x is given by :$A^* A\alpha = A^* Y$, which are normal equations.

$$A = \begin{pmatrix} 1 & x_1 & \cdots & x_1^n \\ 1 & x_2 & \cdots & x_2^n \\ \vdots & \vdots & & \vdots \\ 1 & x_m & \cdots & x_m^n \end{pmatrix}, \quad \alpha = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}$$

- Bisection
- Newton's Iteration

Derived from Taylor's expansion, the form of Newton's Iteration writes:

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

which is of 2-order convergence if x is simple root, and 1-order if x is multiple root in most cases. If x is of multiplicity p, then the form:

$$x_{k+1} = x_k - p\frac{f(x_k)}{f'(x_k)}$$

is of 2-order.

- Secant Method

To avoid use of $f'(x)$, replace newton's method by:

$$x_{k+1} = x_k - \frac{f(x_k)(x_k - x_{k-1})}{f(x_k) - f(x_{k-1})}$$

- Fixed-point iteration

The iteration has the form:

$$x_{k+1} = \varphi(x_k)$$

Definition of Contractive Mapping

A mapping φ is called contractive if there exists $\lambda<1$ such that $|\varphi(x) - \varphi(y)| \leq \lambda|x - y|$ for all x,y in the domain of φ.

Theorem on contractive mapping

Let C be a closed set on real line. If φ is a contractive mapping of C into C, then it has a unique fixed point. And the fixed-point iteration is convergent from any starting point in C.

Remark: This is another way to describe the theorem 3.1 in the textbook.

4.Solution of linear systems of equations

- Gauss Elimination $-O(n^3)$
- LU Decomposition –Find lower triangular L and upper triangular U such that A=LU. (Not unique)

Doolittle: $l_{ii} = 1$. Crout: $u_{ii} = 1$. Cholesky: $U = L^T$

Theorem 1 (sufficient condition)

If all n leading principal minors of the n× n matrix A are nonsingular, then A has an LU-decomposition

Theorem 2

If A is a real, symmetric, and positive definite matrix, then it has a unique factorization, $A = LL^T$, in which L is lower triangular with a positive diagonal.

Remark: Gauss elimination is LU decomposition in essence.

- Diagonally Diminant Matrices:
  $|a_{ii}| > \sum_{j=1,j\neq i}^{n} |a_{ij}|, (1 \leq i \leq n)$

Theorem 3

Gauss elimination without pivoting preserves the diagonal dominance of a matrix.

Corollary

Every diagonally dominant matrix is nonsingular and has an LU-factorization

Vector norms:

1. $\|x\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2}$
2. $\|x\|_\infty = \max_{1 \le i \le n} |x_i|$
3. $\|x\|_1 = \sum_{i=1}^{n} |x_i|$

Matrix norms(Defined after vector norm):

1. $\|A\|_\infty = \max_{1 \le i \le n} \sum_{j=1}^{n} |a_{ij}|$
2. $\|A\|_1 = \max_{1 \le j \le n} \sum_{i=1}^{n} |a_{ij}|$
3. $\|A\|_2 = \sqrt{\rho(A^T A)}$

Condition number:$\kappa(A) = \|A\|\|A^{-1}\|$

Theorem on bounds involving condition number

In solving systems of equations Ax=b, the conditon number $\kappa(A)$, turbulence and solution satisfies:

$$\frac{1}{\kappa(A)} \frac{\|r\|}{\|b\|} \le \frac{\|e\|}{\|x\|} \le \kappa(A) \frac{\|r\|}{\|b\|}$$

- Iteration method

Introduce splitting matrix Q, and the system of linear equations $Ax = b$ can be written as: $Qx = (Q - A)x + b$. If Q is nonsingular, then there is iteration formula:$x^{(k+1)} = (I - Q^{-1}A)x^k + Q^{-1}b$
Generally, an iteration method has the form:$x^{(k+1)} = Gx^{(k)} + c$,

Theorem 4 (Sufficient Condition)

If $\|I - Q^{-1}A\| < 1$ then the iteration converges from any starting point $x^{(0)}$.(Proof)

Theorem 5 (Sufficient and necessary)

In order that the iteration formula:

$$x^{(k+1)} = Gx^{(k)} + c$$

produce a sequence converging to $(I - G)^{-1}c$, for any starting vector $x^{(0)}$, it is necessary and sufficient that $\rho(G) < 1$.

	Splitting mtx	Iteration mtx
Jacobi	D	$I - D^{-1}A = -D^{-1}(L+U)$
Gauss-Seidel	D+L	$-(D+L)^{-1}U$
SOR	$\frac{1}{\omega}D + L$	$(D+\omega L)^{-1}((1-\omega)D - \omega U)$

- SOR Iteration

If we view Gauss-Seidel iteration as: $x^{(k+1)} = x^{(k)} - D^{-1}r^{(k)}$, then we want to modify the item $-D^{-1}r^{(k)}$, so that the formula can converge faster. Introduce ω, SOR iteration has the form:

$$x^{(k+1)} = x^{(k)} - \omega D^{-1}r^{(k)}$$

Jacobi:

$$\begin{cases} x_1^{(k+1)} = -\frac{1}{a_{11}}(a_{12}x_2^{(k)} + \cdots + a_{1n}x_n^{(k)} - b_1) \\ x_2^{(k+1)} = -\frac{1}{a_{22}}(a_{21}x_1^{(k)} + a_{23}x_3^{(k)} + \cdots + a_{2n}x_n^{(k)} - b_2) \\ x_3^{(k+1)} = -\frac{1}{a_{33}}(a_{31}x_1^{(k)} + a_{32}x_2^{(k)} + a_{34}x_4^{(k)} + \cdots + a_{3,n}x_n^{(k)} - b_3) \\ \vdots \\ x_n^{(k+1)} = -\frac{1}{a_{nn}}(a_{n1}x_1^{(k)} + a_{n2}x_2^{(k)} + \cdots + a_{n,n-1}x_{n-1}^{(k)} - b_n) \end{cases}$$

Gauss-Seidel:

$$\begin{cases} x_1^{(k+1)} = -\frac{1}{a_{11}}(a_{12}x_2^{(k)} + \cdots + a_{1n}x_n^{(k)} - b_1) \\ x_2^{(k+1)} = -\frac{1}{a_{22}}(a_{21}x_1^{(k+1)} + a_{23}x_3^{(k)} + \cdots + a_{2n}x_n^{(k)} - b_2) \\ x_3^{(k+1)} = -\frac{1}{a_{33}}(a_{31}x_1^{(k+1)} + a_{32}x_2^{(k+1)} + a_{34}x_4^{(k)} + \cdots + a_{3,n}x_n^{(k)} - b_3) \\ \vdots \\ x_n^{(k+1)} = -\frac{1}{a_{nn}}(a_{n1}x_1^{(k+1)} + a_{n2}x_2^{(k+1)} \cdots + a_{n,n-1}x_{n-1}^{(k+1)} - b_n) \end{cases}$$

SOR:

$$\begin{cases} x_1^{(k+1)} = x_1^{(k)} - \frac{\omega}{a_{11}}(a_{11}x_1^{(k)} + a_{12}x_2^{(k)} + \cdots + a_{1n}x_n^{(k)} - b_1) \\ x_2^{(k+1)} = x_2^{(k)} - \frac{\omega}{a_{22}}(a_{21}x_1^{(k+1)} + a_{22}x_2^{(k)} + a_{23}x_3^{(k)} + \cdots + a_{2n}x_n^{(k)} - b_2) \\ x_3^{(k+1)} = x_3^{(k)} - \frac{\omega}{a_{33}}(a_{31}x_1^{(k+1)} + a_{32}x_2^{(k+1)} + a_{33}x_3^{(k)} + \cdots + a_{3,n}x_n^{(k)} - b_3) \\ \vdots \\ x_n^{(k+1)} = x_n^{(k)} - \frac{\omega}{a_{nn}}(a_{n1}x_1^{(k+1)} + a_{n2}x_2^{(k+1)} \cdots + a_{n,n-1}x_{n-1}^{(k+1)} + a_{nn}x_n^{(k)} - b_n) \end{cases}$$

Theorem on Convergence 1

If A is diagonally dominant, then Jacobi and Gauss-Seidel converges from any starting point.

Theorem on Convergence 2

If A is real, symmetric and positive definite, then SOR converges from any starting point for any $0 < \omega < 2$.(Including Gauss-Seidel)