NA in One Paper

Jin HUANG

1 Error

- Cause: Truncation, Round-off.
- Measurement: Absolute, Relative.
- Propagation: Viewing output error as the function of input error.

2 Main Tools

1. The method of undetermined coefficients: form and coefficients.
2. Series: Taylor series, Geometric series.
3. Norm of vector and matrix, metric, orthogonality.

3 Solving Equation

3.1 $f(x) = 0, x \in \mathbb{R}$

Error $e_n = \|x_n - x\|$.

- Convergence rate: $\lim_{n\to\infty} e_{n+1}/e_n^\alpha = \gamma, \gamma \neq 0$.
- Taylor of fixed-point iteration: $f(x) = 0 \Rightarrow g(x) - x = 0$.

$$\begin{aligned} e_{n+1} &= \|g(x_n) - g(x)\| \\ &= \|g'(\xi_n)(x_n - x) + \cdots\| \leq \|g'(\xi_n)\| e_n + \cdots. \end{aligned}$$

Special notices:

- Multi-root slows down the convergence of Newton's method: solve $\mu(x) = f(x)/f'(x)$.
- Aitken acceleration: linear convergence.

3.2 $Ax = b, x, b \in \mathbb{R}^n$

Direct method Gaussian Elimination, $A = LU$. (Partial and scaled partial) Pivoting comes from the division and multiplication. For tri-diagonal matrix, using the method of undetermined coefficients.

Iteration Let $A = D - L - U$, from $(D - L - U)x = 0x + b$ to

- Jacobian: $Dx = (L + U)x + b$.
- Gauss-Seidel: $(D - L)x = Ux + b$.
- SOR: $(D/\omega - L)x = ((1/\omega - 1)D + U)x + b$.

All in the form of $x = Tx + c$, and the convergence of $\|x - x_k\|$ is related to $\rho(T)$. The error defined from $(A+\delta A)(x+\delta x) = b+\delta b$ is related to the condition number $K = \|A\| \cdot \|A\|^{-1}$:

$$\frac{\|\delta x\|}{\|x\|} \leq \frac{K(A)}{1 - K(A)\frac{\|\delta A\|}{\|A\|}} \left(\frac{\|\delta A\|}{\|A\|} + \frac{\|\delta b\|}{\|b\|} \right).$$

Approximate eigenvalues Power method and Inverse Power Method: Using eigendecomposition to write matrix polynomial.

$$A^n + pI = U^{-1}(\Lambda^n + pI)U. \tag{1}$$

The largest eigenvalue (and eigenvector) dominate the sequence.

4 Interpolation

A special type of approximation with zero error at interpolation points.

All use the method of undetermined coefficients to represent a polynomial in different bases, but have *special efficient ways* to solve the coefficients.

- Power basis: $x^0, x^1, x^2, \cdots$, coefficient: $a_0, a_1, \cdots$.
- Lagrange basis: $L_{n,i}(x) = \Pi_{j=0,i\neq j}^n \frac{x - x_j}{x_i - x_j}$, coefficient: $f(x_i)$. Extending into Hermite method with more equations (e.g. f').
- Newton basis: $N_0 = 1, N_i = \Pi_{j=0}^{j<i}(x - x_j)$, coefficient: $f[x_0, x_1, \cdots, x_i]$.

Remainder from Taylor's expansion: $R_n(x) = \frac{f^{n+1}(\xi_x)}{(n+1)!}\Pi_{i=0}^n (x - x_i)$.

Spline: piece-wise interpolation, form and its coefficients coming from integrating f'' into f' and then f.

5 Approximate

Cannot exact interpolate the points: has error on those points.

5.1 Least square error

The form: linear combination of a set of bases ϕ_i by coefficients c_i.

$$P(x, c) = \sum_i c_i \phi_i(x) = (\phi_i(x)) \cdot c. \tag{2}$$

Solving an optimization about the undetermined coefficients $\{c_i\}$, e.g. an equation about gradient of $c = \{c_i\}$.

- Discrete: $\min_c \sum_k w_k (P(x_k, c) - y_k)^2$.
- Continuous: $\min_c \int_a^b w(x)(P(x, c) - y(x))^2 dx$.

The above optimization involves inner product (with metric w) between ϕ_i, ϕ_j, and thus we discuss orthogonal bases (polynomials) for efficient solving. The method to achieve orthogonal is kind of GramSchmidt orthogonalization.

5.2 Minimize maximal error

Starting from the error at any x

$$\begin{aligned} |P_n(x) - f(x)| &= \left| \frac{f^{(n+1)}(\xi_x)}{(n+1)!} \Pi_{i=0}^{n} (x - x_i) \right| \\ &\triangleq \left| \frac{f^{(n+1)(\xi_x)}}{(n+1)!} \right| |\Pi_{i=0}^{n} (x - x_i)| \end{aligned} \tag{3}$$

we are asking to minimize $\max |\Pi_{i=0}^{n}(x - x_i)|$. The idea is to make the polynomial $T_n(x) = \Pi_{i=0}^{n}(x - x_i)$ vibrate in a narrowest band: equal amplitude vibration. Turn $\cos(nx)$ into the polynomials version cos, i.e. Chebyshev polynomial

$$T_n(x) = \cos(n \arccos(x)), x \in [-1, 1]. \tag{4}$$

Then using the roots of T_n in Lagrange interpolation for P_n. Compare to many previous interpolation/approximation method, we "optimize" basis ϕ in a constructive way. In other words, the undetermined coefficients include the bases ϕ instead of just c.

Economization of Power Series: elimination using T_n.

6 Numerical Differentiation and Integration

Error of Lagrange interpolation/approximation:

$$\begin{aligned} f(x) &= \sum_{k=0}^{n} f(x_k) L_k(x) + \frac{f^{(n+1)}(\xi_x)}{(n+1)!} \Pi_{i=0}^{n} (x - x_i) \\ &\triangleq P_n(x) + R_n(x). \end{aligned} \tag{5}$$

6.1 Differentiation

Differentiate eq. (5).

Another important way to think about this problem: using the method of undetermined coefficients to eliminate irrelevant terms in Taylor series at points around x.

6.2 Integration

Integrate eq. (5). Integration can have zero error even $R_n(x) \neq 0$, because of the integration. The accuracy is defined as the largest n that $\int_a^b R_n(x) = 0$. n is always odd.

Composite Similar to the idea of spline, piece-wise low order.

Romberg integration Similar to Aitken, use more equations coming from interval subdivision to improve accuracy. Look the error respect to the interval.

Richardson extrapolation The same to the above, but more general.

Adaptive Only pay the cost at necessary region for economic computation. The key is to "estimate the error of region: using an attempt subdivide and check the relationship between the two errors.

Gaussian quadrature Similar to Chebyshev polynomial, optimize the basis (integration points), or the undetermined coefficients include the bases ϕ instead of just c. To solve the nonlinear optimization efficiently, we ask $\Pi_{i=0}^{n}(x - x_i)$ orthogonal (under specific metric) to all the lower degree polynomials (read the proof), and use its root x_i to construct ϕ, and then simply solve c.

7 IVP for ODE

Using a set of methods with the same order of local truncation error.

- Single step explicit method for *the first* m initial values.
- multiple step explicit method.
- multiple step implicit method.

Basic ideas: Taylor expansion or integration.

7.1 Single step

Taylor's method Use the Taylor's expansion of $y'(t)$, i.e. $f(t, y)$ with respect to t. Euler's method is kind of Taylor's method.

Runge-Kutta method Apply the method of undetermined coefficients to Taylor's expansion: recursively on many points in $[0, h]$.

7.2 Multiple step

Viewpoint of integration: Newton interpolation on f and then integrate.

Viewpoint of Taylor's expansion: taking y_{i+1}'s Taylor's expansion as ground truth, apply the method of undetermined coefficients to composite Taylor's expansion of w_i, f_i.

7.3 High order and Systems of ODE

Turn high order ODE into Systems of 1^{st}-order ODE: a vector form ODE.

7.4 Stability

Local truncation error considers step size h. Stability considers error propagation with respect to t.

The idea: using test equation $y' = \lambda y, y(0) = \alpha$, inject error ε into $y(0)$, then check the error at $y(t)$. Usually get a geometric series about $H = \lambda h$. For system, use the eigendecomposition to decouple the variables, and then get a set of independent ODEs.