《数值分析》第九章

常微分方程初值问题的数值解法

包承龙

YMSC

2019 年 12 月 15 日

目录

目录

常微分方程初值问题

许多常见、重要的自然科学和人文科学及工程技术中许多问题都可以用常微分方程 (组) 初值问题来描述. 通常来说这些问题不能写出其解的解析表达式，因此需要用数值方法求解.

考虑如下一阶非线性常微分方程初值问题

$$\begin{cases} \frac{dy}{dt} = f(t, y) \\ y(0) = y_0 \end{cases} \tag{1}$$

若满足

$$|f(x, y_1) - f(x, y_2)| \leqslant L|y_1 - y_2| \qquad (x, y_1), (x, y_2) \in D \tag{2}$$

则称 f 关于变量 y 满足 Lipschitz 条件，L 称为 f 的 Lipschitz 常数.

常微分方程初值问题

适定性

如果初值问题 (2) 满足

- 存在唯一解；
- 对于任给的 $\epsilon > 0$，存在正数 $\delta^\star$，使得当 $|\epsilon_0| < \delta^\star$ 时，在 $[x_0, b]$ 上 δ 连续，且 $||\delta|| < \delta^\star$ 时，初值问题

$$\begin{cases} z_x = f(x, z) + \delta, & x \in [x_0, b] \\ z(0) = z_0 + \varepsilon_0 \end{cases} \tag{3}$$

存在唯一解，且 $\forall x \in [x_0, b]$，都有 $||z - y|| < \epsilon$，则称初值问题 (2) 是适定的.

定理

设 $D = [x_0, b] \times \mathbb{R}$，$f$ 在 D 上连续并满足 Lipschitz 条件，那么初值问题 (2) 是适定的.

目录

Euler 方法求解方程

若已知初值 $y(t_0)=y_0$，若 $y(t)$ 充分光滑，由 Tayler 展开

$$y(t_1)=y(t_0+h)=y(t_0)+hy'(t_0)+\frac{h^2}{2}y''(\xi),\quad \xi\in(t_0,t_1) \tag{4}$$

舍去高阶项可以得到 $y(t_1)\approx y_0+hf(t_0,y_0)$.
有了 y_0 后可以继续求解下去，得到

$$\begin{cases} y_{n+1}=y_n+hf(t_n,y_n) \\ y_0=y(t_0) \end{cases} \tag{5}$$

此即 Euler 方法. 当 $y(t_0)$ 已知时此为显式方法，且该迭代步骤是单步的，被称为显式单步法.

Euler 方法求解方程

还可以从积分形式来推导 Euler 方法:
由初值问题的等价形式

$$y(t) = y_0 + \int_{t_0}^{t} f(s, y(s))ds \tag{6}$$

可得

$$y(t_1) = y_0 + \int_{t_0}^{t_1} f(s, y(s))ds \tag{7}$$

之后若用左侧积分公式计算可得

$$y_1 = y_0 + hf(t_0, y_0) \tag{8}$$

即 Euler 方法.

Euler 方法求解方程——隐式方法

若在式 (7) 中使用右侧积分公式，可以得到

$$y_{n+1} = y_n + hf(t_{n+1}, y_{n+1}) \tag{9}$$

该方法被称为后退 Euler 方法或隐式 Euler 法. 此时如果 $f(t, y)$ 为关于 y 的非线性函数，则该隐式 Euler 法需要迭代求解. 可以先用显式 Euler 法给出初值，然后使用 Picard 迭代

$$\begin{aligned} y_{n+1}^{(0)} &= y_n + hf(t_n, y_n) \\ y_{n+1}^{(1)} &= y_n + hf\left(t_n, y_{n+1}^{(0)}\right) \\ &\cdots\cdots \\ y_{n+1}^{(k+1)} &= y_n + hf\left(t_n, y_{n+1}^{(k)}\right) \end{aligned} \tag{10}$$

当 $\left|y_{n+1}^{(k+1)} - y_{n+1}^{(k)}\right| < \varepsilon$ 时停止迭代，输出 $y_{n+1} = y_{n+1}^{(k+1)}$.

Euler 方法求解方程——隐式方法

对该迭代法，若假设 $f(t, y)$ 关于 y 满足 Lipschitz 条件

$$|f(t, y_1) - f(t, y_2)| \leq L\,|y_1 - y_2| \tag{11}$$

则有估计

$$\begin{aligned}\left|y_{n+1} - y_{n+1}^{(k+1)}\right| &= h\left|f(t_{n+1}, y_{n+1}) - f\left(t_n, y_{n+1}^{(k)}\right)\right| \\ &\leq \mathrm{Lh}\left|y_{n+1} - y_{n+1}^{(k)}\right| \leq \cdots \\ &\leq (Lh)^{k+1}\left|y_{n+1} - y_{n+1}^{(0)}\right|\end{aligned} \tag{12}$$

这说明只要 $Lh < 1$，该迭代法就会收敛.

Euler 方法求解方程——梯形方法与改进 Euler 法

对于积分 (7) 如果使用梯形公式计算

$$y_{n+1} = y_n + \frac{h}{2}\left[f(t_n, y_n) + f(t_{n+1}, y_{n+1})\right] \qquad n \geq 0 \tag{13}$$

该方法被称为梯形法，其也是一种隐式格式，可以通过迭代的方式求解，其 Picard 迭代公式为

$$\begin{cases} y_{n+1}^{(0)} = y_n + hf(t_n, y_n) \\ y_{n+1}^{(k+1)} = y_n + \dfrac{h}{2}\left[f(t_n, y_n) + f\left(t_{n+1}, y_{n+1}^{(k)}\right)\right] \qquad k = 0, 1, \cdots \end{cases} \tag{14}$$

可以证明，其迭代收敛条件为 $Lh/2 < 1$.

Euler 方法求解方程——梯形方法与改进 Euler 法

以上每次计算都要计算非线性函数 $f(t, y)$ 的值，如果每一步时间步都计算多次的话，这个计算量会较大.
而如果我们用 Euler 法提供初值的一个预估之后，在进行迭代校正，这将加快计算速度. 该方法被称为改进的 Euler 法. 其迭代步骤为

$$\begin{cases} \bar{y}_{n+1} = y_n + hf(t_n, y_n) \\ y_{n+1} = y_n + \frac{h}{2}\left[f(t_n, y_n) + f(t_{n+1}, \bar{y}_{n+1})\right] \end{cases} \tag{15}$$

将一式带入二式可以得到

$$y_{n+1} = y_n + \frac{h}{2}\left[f(t_n, y_n) + f(t_{n+1}, y_n + hf(t_n, y_n))\right] \tag{16}$$

可以看到，该方法为一种显式单步法.

目录

显式单步法

对之前提到的改进 Euler 法

$$y_{n+1} = y_n + \frac{h}{2}\left[f(t_n, y_n) + f(t_{n+1}, y_n + hf(t_n, y_n))\right] \tag{17}$$

如果将其表示为 $y_{n+1} = y_n + h\varphi(x_n, y_n; h)$ 的形式，则

$$\varphi(x, y; h) = \frac{1}{2}\left[f(x, y) + d(x + h, y + hf(x, y))\right] \tag{18}$$

这里 $\varphi(x, y; h)$ 被称为增量函数.

对迭代法算出的值 $\{y_n\}$，其与真值的差 $e_n = y(x_n) - y_n$ 被称为整体截断误差；而若假设前 n 步都精确的情况下，计算一步的误差

$$T(x, y; h) = y(x + h) - y(x) - h\varphi(x, y(x); h) \tag{19}$$

$T_{n+1} = t(x_n, y(t_n); h)$ 被称为局部截断误差.

显式单步法——相容性

相容性

如果

$$\lim_{h\to 0}\frac{1}{h}T[t,y(t);h]=0 \tag{20}$$

即差分格式收敛到常微分方程，则称该格式与原格式相容. 而若存在常数 k 使得

$$|T(x,y;h)|\leqslant kh^{p+1} \tag{21}$$

则称其为 p 阶相容格式.

定理

单步法 $y_{n+1}=y_n+h\varphi\left(t_n,y_n;h\right)$ 相容的充分必要条件是

$$\lim_{h\to 0}\varphi(t,y;h)=f(t,y) \tag{22}$$

显式单步法——收敛性

收敛性

设 $T = Nh$，$t_n = nh$，令

$$E(h) = \max_{0 \le n \le N} |\varepsilon_n| \equiv \max_{0 \le n \le N} |y(t_n) - y_n| \tag{23}$$

如果 $\lim_{h\to 0} E(h) = 0$，则称单步法是收敛的. 如果 $\forall h \in (0, h_0]$，都存在常数 $C > 0$s.t.$E(h) < Ch^p$，则称单步法是 p 阶收敛的.

收敛性与相容性的关系

设单步法 $y_{n+1} = y_n + h\varphi(t_n, y_n; h)$ 的增量函数 $\varphi(x, y; h)$ 是连续函数，且关于 y 满足 Lipschitz 条件，即

$$|\phi(t, y_1; h) - \phi(t, y_2; h)| \leqslant L|y_1 - y_2| \tag{24}$$

则单步法收敛的充分必要条件是该单步法是相容的.

Runge-Kutta 方法

Runge-Kutta 方法是通过不同点上函数值的不同组合来提高方法的精度，其基本形式为

$$y_{n+1} = y_n + h\varphi(x_n, y_n; h) \tag{25}$$

其中

$$\varphi(x, y; h) = \sum_{i=1}^{s} b_i k_i \tag{26}$$

$$k_i = f(x_n + c_i h, y_n + h\sum_{j=1}^{i} a_{ij} k_j) \tag{27}$$

这里设 $c_i = \sum_{j=1}^{i} a_{ij}$ 为行和条件.

Runge-Kutta 方法——待定系数法

可以通过待定系数法求解.
$s=1$ 时，其为最简单的 Ruler 法 $y_{n+1}=y_n+hf(t_n,y_n)$；
$s=2$ 时，有条件 $c_1+c_2=1$ 及 $a_2c_2=1/2$. 其自然有无穷多组解，最常用的为以下两种格式：

- 显示中点格式：$c_1=0$，$c_2=1$，$a_2=1/2$，此时

$$y_{n+1}=y_n+hf\left(t_n+\frac{h}{2},y_n+\frac{h}{2}f_n\right) \tag{28}$$

- 改进 Euler 法：$c_1=c_2=1/2$，$a_2=1$，此时

$$y_{n+1}=y_n+\frac{h}{2}\left[f_n+f(t_n+h,y_n+hf_n)\right] \tag{29}$$

Runge-Kutta 方法——待定系数法

$s=4$ 时，由显式条件及相容性

$$a_1=0, a_2=b_{21}, a_3=b_{31}+b_{32}, a_4=b_{41}+b_{42}+b_{43} \tag{30}$$

再比较 φ 和 $\tilde{\varphi}$ 的值，得到一组公式. 其也有无穷多组解，最常用的为一组特殊值

$$\begin{gathered}\mathbf{a}=\left(0,\frac{1}{2},\frac{1}{2},1\right),\quad \mathbf{c}=\left(\frac{1}{6},\frac{1}{3},\frac{1}{3},\frac{1}{6}\right)\\ b_{21}=\frac{1}{2},b_{31}=0,\quad b_{32}=\frac{1}{2},\quad b_{41}=0=b_{42},\quad b_{43}=1\end{gathered} \tag{31}$$

该方法为经典的 4 阶 Runge-Kutta 方法.

Runge-Kutta 方法——隐式法

也可以采用隐式的 Runge-Kutta 法，对之前的积分 $y(s_i) = y(t_n) + \int_{t_n}^{s_i} f(t, u(t))dt$ 利用 s_j 上的值进行估计，得到

$$y(s_i) = y(t_n + a_i h) \approx y_n + h\sum_{i=1}^{s} b_{ij} f(s_j, y(s_j)) \tag{32}$$

将 K_j 带入即得到隐式 Runge-Kutta 方法

$$\begin{cases} y_{n+1} = y_n + \sum_{i=1}^{s} c_i K_i \\ K_i = f\left(t_n + ha_i, y_n + h\sum_{j=1}^{s} b_{ij} K_j\right) \end{cases} \tag{33}$$

隐式法的缺点在于无法显式计算上面的 K_j，通常需要求解一个非线性方程组才能得到.

目录

线性多步法

在单步法中我们利用插值来计算 K_j，而若我们已经知道了多个节点的值 y_j，$j = 1, \cdots, k-1$ 后，我们可以利用现有的 e 函数进行插值，而不需要引入新的节点，该方法被称为线性多步法.

多步法的一般形式可以表示为

$$\sum_{j=0}^{k} \alpha_j y_{n+j} = h \sum_{j=0}^{k} \beta_j f(t_{n+j}, y_{n+j}) \tag{34}$$

这里一般设 $\beta_k \neq 0$，因若 $\beta_k = 0$ 则其退化为显示格式.

线性多步法

对于多步法，同样地可以计算其局部截断误差

$$T[y(t_n);h]=\sum_{j=0}^{k}[\alpha_j y(t_{n+j})-h\beta_j f(t_{n+j},y(t_{n+j}))] \tag{35}$$

将其在 t_n 处做 Taylor 展开，可以得到

$$T[y(t_n);h]=c_0 y(t_n)+c_1 h y'(t_n)+\cdots+c_p h^p y^{(p)}(t_n)+\cdots \tag{36}$$

可以得到

$$\begin{cases} c_0=\alpha_0+\alpha_1+\cdots+\alpha_k \\ c_1=\alpha_1+2\alpha_2+\cdots+k\alpha_k-(\beta_0+\beta_1+\cdots+\beta_k) \\ \cdots\cdots \\ c_p=\frac{\alpha_1+2^p\alpha_2+\cdots+k^p\alpha_k}{p!}-\frac{\beta_1+2^{p-1}\beta_2+\cdots+k^{p-1}\beta_k}{(p-1)!} \\ \cdots\cdots \end{cases} \tag{37}$$

线性多步法

这里若 $c_0 = \cdots = c_p = 0$，$c_{p+1} \neq 0$，则称该多步法为 p 阶方法，其局部截断误差为

$$T[y(t_n);h] = c_{p+1}h^{p+1}y^{(p+1)}(\xi_n) \tag{38}$$

线性多步法——数值积分

也可以通过数值积分的方法进行推导. 将常微分方程 (3) 在 $[t_{n+k-1}, t_{n+k}]$ 上积分得到

$$y(t_{n+k}) = y(t_{n+k-l}) + \int_{t_{n+k-l}}^{t_{n+k}} f(t, y(t))\,dt \tag{39}$$

然后通过插值节点的各种插值多项式代替被积函数 $f(t, y(t))$ 可以得到各种各样的线性多步法.

若取 $k = l = 4$，得到

$$y(t_{n+4}) = y(t_n) + \int_{t_n}^{t_{n+4}} f(t, y(t))\,dt \tag{40}$$

并取 $f(t, y(t))$ 在 t_{n+1}，t_{n+2}，t_{n+3} 上的值进行二次插值，得到

$$L_2(t) = \sum_{j=0}^{2} f(t_{n+1+j}, y(t_{n+1+j}))\, l_j(t) \tag{41}$$

线性多步法——Milne 方法

将其代替 $f(t, y(t))$ 并积分可以得到

$$\int_{t_n}^{t_{n+4}} L_2(t)dt = \frac{4h}{3}\left[2f_{n+1} - f_{n+2} + 2f_{n+3}\right], \quad \dot{\mathfrak{x}}\mathbb{Z}, f_m = f(t_m, y_m) \tag{42}$$

因而得到格式

$$y_{n+4} = y_n + \frac{4h}{3}\left[2f_{n+1} - f_{n+2} + 2f_{n+3}\right] \tag{43}$$

该格式被称为 Milne 方法，其局部截断误差为

$$T_{n+4} = \frac{14}{45}h^5 y^{(5)}\left(t_n\right) + \mathcal{O}\left(h^6\right) \tag{44}$$

因而其为四阶方法.

线性多步法——Adams 方法

如果取 $l=1$，得到

$$y(t_{n+k}) = y(t_{n+k-1}) + \int_{t_{n+k-1}}^{t_{n+k}} f(t, y(t))dt \tag{45}$$

如果利用 $f(t_{n-k+1}), \cdots, f(t_n)$ 上的值构造 $k-1$ 阶多项式 $L_{k-1}(t)$ 代替 f，可以得到 Adams 外插格式

$$y_{n+k} = y_{n+k-1} + h\sum_{j=0}^{k-1} \beta_j f_{n+j} \tag{46}$$

而如果利用 $f(t_{n+k}), \cdots, f(t_n)$ 上的值构造 k 阶多项式 $\widetilde{L}_k(t)$ 代替 f。可以得到 Adams 内插格式

$$y_{n+k} = y_{n+k-1} + h\sum_{j=0}^{k} \widetilde{\beta}_j f_{n+j} \tag{47}$$

目录

线性多步法的相容性

利用截断误差的定义，如果

$$T[y(t_n);h] = \sum_{j=0}^{k} [\alpha_j y(t_{n+j}) - h\beta_j f(t_{n+j}, y(t_{n+j}))] = c_{p+1} h^{p+1} y^{(p+1)}(\xi_n) \tag{48}$$

则称多步法是 p 阶相容的.

如果引入第一、第二特征多项式

$$\rho(\lambda) = \sum_{j=0}^{k} \alpha_j \lambda^j, \quad \sigma(\lambda) = \sum_{j=0}^{k} \beta_j \lambda^j \tag{49}$$

可以证明，多步法 (34) 与 (3) 相容，当且仅当 $\rho(1) = 0$，且 $\rho'(1) = 1$.

线性多步法的收敛性

收敛性

称多步法 (34) 是收敛的，如果存在不依赖于 h 的常数 C 及 $h_0, \varepsilon > 0$，s.t.$\forall h \in (0, h_0)$ 及任两组不同初值得到的解 $\{u_i\}$ 和 $\{v_n\}$，当 $\max_{0\le j\le k-1} |u_j - v_j| < \varepsilon$ 时，就有

$$\max_{nh\le T} |u_n - v_n| \le C \max_{0\le j\le k-1} |u_j - v_j| \tag{50}$$

线性多步法的收敛性

Gronwall 不等式

设 $|\xi_j|$ 是 $\mathbb{R}$ 中数列并满足，对于常数 $A > 0$ 和 $B \geqslant 0$ 有

$$|\xi_j| \leqslant A \sum_{m=0}^{j-1} |\xi_m| + B \qquad j = 1, 2, \cdots \tag{51}$$

则有不等式

$$|\xi_j| \leqslant (A|\xi_0| + B) e^{(j-1)A} \qquad j = 1, 2, \cdots \tag{52}$$

利用 Gronwall 不等式可以证明

多步法收敛的充要条件

多步法 (34) 收敛，当且仅当第一多项式 $\rho(\lambda)$ 满足根条件：$\rho(\lambda)$ 的根都在单位圆内，且在单位圆周上的根全为单根.

线性多步法的稳定性

绝对稳定性

称线性多步法 (34) 关于 $\bar{h} = \mu h$ 是绝对稳定的，如果特征方程的根都在单位圆周内，即

$$\lambda_j(\bar{h}) < 1 \qquad j = 1, \cdots, k \tag{53}$$

绝对稳定性的必要条件

若多步法 (34) 是相容且对初值稳定的格式，且 $h \to 0$ 时关于 $\bar{h} = \mu h$ 绝对稳定，则必有 $Re(\mu) < 0$.

相对绝对稳定性

称多步法 (34) 是相对稳定的，如果特征方程的根满足 $\lambda_1(\bar{h}) = e^{\bar{h}} + \mathcal{O}\left(h^{p+1}\right)$，且

$$\left|\lambda_1(\bar{h})\right| > \left|\lambda_j(\bar{h})\right| \qquad j = 2, \cdots, k \tag{54}$$