

# 常微分方程的数值解法

## well-posed problem

- well-posed :
  - 该常微分方程在给定区间上有唯一解。
  - 该微分方程对小扰动稳定。
- 一个常微分方程

$$y' = f(t, y), \quad a \leq t \leq b \quad -\infty \leq y \leq \infty$$

是 well-posed 当且仅当  $f(t, y)$  对  $y$  满足 lipschitz 条件, 且  $f(t, y)$  在给定区间上连续。

## 局部截断误差

- 定义：对于差分法

$$w_0 = \alpha$$

$$w_{i+1} = w_i + h\phi(t_i, w_i), \quad i = 0, 1, \dots, N-1$$

局部截断误差定义为：

$$\tau_{i+1}(h) = \frac{y_{i+1} - (y_i + h\phi(t_i, y_i))}{h} = \frac{y_{i+1} - y_i}{h} - \phi(t_i, y_i)$$

## 欧拉方法

- 基本思想：  
由泰勒公式

$$y(t_{i+1}) = y(t_i) + hy'(t_i) + \frac{h^2}{2}y''(\xi)$$

这里  $t_i$  等距离取点,  $h$  表示步长。考虑到微分方程有  $y' = f(t, y)$ , 进一步有

$$y(t_{i+1}) = y(t_i) + hf(t_i, y(t_i)) + \frac{h^2}{2}y''(\xi)$$

去掉尾项, 并令  $w_i = y(t_i)$  则有

$$w_{i+1} = w_i + hf(t_i, w_i)$$

以序列  $\{w_i\}$  近似代替  $\{y(t_i)\}$  即得最终答案。

- 误差：

$$|y(t_i) - w_i| \leq \frac{hM}{2L} [e^{L(t_i-a)} - 1]$$

随着步长线性减小, 若考虑扰动,

$$|y(t_i) - w_i| \leq \frac{1}{L} \left( \frac{hM}{2} + \frac{\delta}{h} \right) [e^{L(t_i-a)} - 1] + \delta_0 e^{L(t_i-a)}$$

## 高阶泰勒方法

- 基本思想：欧拉方法只是用了泰勒展开式中的第一项, 试图用展开式中更多项来提高精度。
- 基本过程：  
将  $y(t)$  在  $t_i$  处展开到  $n$  阶, 有

$$y(t_{i+1}) = y(t_i) + hy'(t_i) + \frac{h^2}{2}y''(t_i) + \cdots + \frac{h^n}{n!}y^{(n)}(t_i) + \frac{h^{n+1}}{(n+1)!}y^{(n+1)}(\xi_i)$$

考虑到  $y'(t) = f(t, y)$ ，带入后得到：

$$y(t_{i+1}) = y(t_i) + hf(t_i, y(t_i)) + \frac{h^2}{2}f'(t_i, y(t_i)) + \cdots + \frac{h^n}{n!}f^{(n-1)}(t_i, y(t_i)) + \frac{h^{n+1}}{(n+1)!}f^{(n)}(\xi_i, y(\xi_i))$$

引入记号  $T^n(t, y)$  将表达式简写作：

$$y(t_{i+1}) = y(t_i) + hT^n(t_i, y_i) + \frac{h^{n+1}}{(n+1)!}f^{(n)}(\xi_i, y(\xi_i))$$

- 局部截断误差：对于 n 阶的泰勒方法，如果采用步长 h 并且  $y(t)$  是 n+1 阶连续的，那么局部阶段误差是  $O(h^n)$ 
  - 从泰勒多项式的余项和局部截断误差的定义立得。

## Runge-Kutta 方法

- 基本思想：在 n 阶泰勒方法中，需要求解函数的直到 n 阶的导数，这个工作量很大，试图采用别的方法在取得同量级精度的同时能减少计算量。
- 二阶 Runge-Kutta 方法：
  - 试图用  $a_1f(t + \alpha_1, y + \beta_1)$  近似表示  $T^2(t, y) = f(t, y) + \frac{h}{2}f'(t, y)$
  - 利用二元函数的求偏导数：

$$f'(t, y) = \frac{df}{dt}(t, y) = \frac{\partial f}{\partial t}(t, y) + \frac{\partial f}{\partial y}(t, y) \cdot y'(t)$$

所以有

$$T^{(2)}(t, y) = f(t, y) + \frac{h}{2}\frac{\partial f}{\partial t}(t, y) + \frac{h}{2}\frac{\partial f}{\partial y}(t, y) \cdot f(t, y)$$

再利用二元函数泰勒展开式

$$a_1f(t + \alpha_i, y + \beta_i) = a_1f(t, y) + a_1\alpha_1\frac{\partial f}{\partial t}(t, y) + a_1\beta_1\frac{\partial f}{\partial y}(t, y) + a_1 \cdot R_1(t + a_1, y + \beta_1)$$

尝试调整参数使两者近似相等，得到

$$a_1 = 1, \alpha_1 = \frac{h}{2}, \beta_1 = \frac{h}{2}f(t, y)$$

于是得到

$$T^{(2)}(t, y) = f(t + \frac{h}{2}, y + \frac{h}{2}f(t, y)) - R_1(t + \frac{h}{2}, y + \frac{h}{2}f(t, y))$$

由二元函数泰勒展开式的余项公式可知：

$$R_1(t + \frac{h}{2}, y + \frac{h}{2}f(t, y)) = \frac{h^2}{8}\frac{\partial^2 f}{\partial t^2}(\xi, \mu) + \frac{h^2}{4}f(t, y)\frac{\partial^2 f}{\partial t \partial y}(\xi, \mu) + \frac{h^2}{8}(f(t, y))^2\frac{\partial^2 f}{\partial y^2}(\xi, \mu)$$

于是当所有二阶偏导数都有界时，用这个估计同样可以得到 2 阶的局部截断误差。

- 最终结果：中点公式

$$w_0 = \alpha$$

$$w_{i+1} = w_i + hf(t_i + \frac{h}{2}, w_i + \frac{h}{2}f(t_i, w_i)), \text{ for } i = 0, 1, \dots, N-1$$

- 改进的欧拉方法

$$w_0 = \alpha$$

$$w_{i+1} = w_i + \frac{h}{2}[f(t_i, w_i) + f(t_{i+1}, w_i + hf(t_i, w_i))]$$

- 四阶 Runge-Kutta 方法:

$$\begin{aligned}
 w_0 &= \alpha \\
 k_1 &= hf(t_i, w_i) \\
 k_2 &= hf(t_i + \frac{h}{2}, w_i + \frac{1}{2}k_1) \\
 k_3 &= hf(t_i + \frac{h}{2}, w_i + \frac{1}{2}k_2) \\
 k_4 &= hf(t_{i+1}, w_i + k_3) \\
 w_{i+1} &= w_i + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4)
 \end{aligned}$$

- 主要开销: 计算函数值, 但不用计算导数

# Runge-Kutta-Fehlberg 方法

- 基本思想: 自动调整步长的 Runge-Kutta 方法。利用两种计算方式的差值近似等于局部截断误差, 通过调整步长将局部截断误差限制在一定范围内。
- 基本方法: 选取具有5阶局部阶段误差的 Runge-Kutta 方法

$$\tilde{w}_{i+1} = w_i + \frac{16}{135}k_1 + \frac{6656}{12825}k_3 + \frac{28561}{56430}k_4 - \frac{9}{50}k_5 + \frac{2}{55}k_6$$

和具有4阶局部截断误差的 Runge-Kutta 方法

$$w_{i+1} = w_i + \frac{25}{216}k_1 + \frac{1408}{2565}k_3 + \frac{2197}{4104}k_4 - \frac{1}{5}k_5$$

其中

$$\begin{aligned}
 k_1 &= hf(t_i, w_i) \\
 k_2 &= hf(t_i + \frac{h}{4}, w_i + \frac{1}{4}k_1) \\
 k_3 &= hf(t_i + \frac{3}{8}h, w_i + \frac{3}{32}k_1 + \frac{9}{32}k_2) \\
 k_4 &= hf(t_i + \frac{12}{13}h, w_i + \frac{1932}{2197}k_1 - \frac{7200}{2197}k_2 + \frac{7296}{2197}k_3) \\
 k_5 &= hf(t_i + h, w_i + \frac{439}{216}k_1 - 8k_2 + \frac{3680}{513}k_3) - \frac{845}{4104}k_4 \\
 k_6 &= hf(t_i + \frac{h}{2}, w_i - \frac{8}{27}k_1 + 2k_2 - \frac{3544}{2565}k_3 + \frac{1859}{4104}k_4 - \frac{11}{40}k_5)
 \end{aligned}$$

用  $|\tilde{w}_{i+1} - w_{i+1}|$  作为  $|y(t_{i+1}) - w_{i+1}|$  的估计值。如果误差大于误差限, 则调整步长  $h$  为  $qh$  其中

$$q \leq (\frac{h\epsilon}{|\tilde{w}_{i+1} - w_{i+1}|})^{\frac{1}{n}}$$

同时重新计算这一步的预计值  $w'(t_{i+1})$ , 如果没有超过误差限, 则只调整步长, 不重新计算这一次的值, 直接进入下一次计算。

# 多步方法

- 基本思想: 前面用到的方法只用到了一个点  $t_i$  的信息来估计  $t_{i+1}$  的值, 又知道误差是随着计算次数不断增大的, 试图利用更靠前也更精确的数据来进行估计, 来提高精度。
- 定义:

$$\begin{aligned}
 w_{i+1} &= a_{m-1}w_i + a_{m-2}w_{i-1} + \cdots + a_0w_{i+1-m} + \\
 &h(b_m f(t_{i+1}, w_{i+1}) + b_{m-1}f(t_i, w_i) + \cdots + b_0f(t_{i+1-m}, w_{i+1-m}))
 \end{aligned}$$

- 开（显式）方法和闭（隐式）方法: 如果  $w_{i+1}$  出现在了右端, 则称为隐式方法, 因为在迭代前需要解方程将  $w_{i+1}$  显式的表达出来; 如果  $w_{i+1}$  不出现在右端, 则称为显式方法; 一般来说, 隐式方法闭显式方法要更加精确。
- 经典的多步方法:

- 四阶 Adams-Moulton technique (隐式)

$$w_0 = \alpha, \quad w_1 = \alpha_1, \quad w_2 = \alpha_2$$

$$w_{i+1} = w_i + \frac{h}{24}(9f(t_{i+1}, w_{i+1}) + 19f(t_i, w_i) - 5f(t_{i-1}, w_{i-1}) + f(t_{i-2}, w_{i-2}))$$

- 四阶 Adams-Bashforth technique (显式)

$$w_0 = \alpha, \quad w_1 = \alpha_1, \quad w_2 = \alpha_2, \quad w_3 = \alpha_3$$

$$w_{i+1} = w_i + \frac{h}{24}(55f(t_i, w_i) - 59f(t_{i-1}, w_{i-1}) + 37f(t_{i-2}, w_{i-2}) - 9f(t_{i-3}, w_{i-3}))$$

- 初始时的几个取值是通过单步方法（一般是4阶 Runge-Kutta 方法）求得的。

- 推导方式：通过对插值函数在一定区间上积分得到。

- 局部截断误差：

$$\tau_{i+1}(h) = \frac{y(t_{i+1}) - a_{m-1}y(t_i) - \cdots - a_0y(t_{i+1-m})}{[b_m f(t_{i+1}, y(t_{i+1})) + \cdots + b_0 f(t_{i+1-m}, y(t_{i+1-m}))]} - \frac{h}{b_m f(t_{i+1}, y(t_{i+1})) + \cdots + b_0 f(t_{i+1-m}, y(t_{i+1-m}))}]$$

- 四阶 Adams-Moulton technique (隐式) 有五阶误差
- 四阶 Adams-Bashforth technique (显式) 有四阶误差

## Predictor-Corrector Methods

- 基本思想：隐式方法的误差阶数更高，但解方程的过程很复杂；显式方法的误差阶数低，但是计算简单。考虑将两者结合起来，先用显式方法求出  $w_{i+1}$  的估计值，再带入隐式方法的方程左端，重新计算  $w_{i+1}$ ，进行“纠正”。

NOTE：可以考虑到是否能反复带入隐式方法的表达式来获得更加精确的解？  
一般不采用，反复迭代后趋向的是隐式方程的解，不是真实解。一般通过减小步长来获得更加精确的解。

- 一般采用四阶 Adams-Bashforth technique (显式) 作为 Predictor，四阶 Adams-Moulton technique (隐式) 作为 Corrector，用四阶 Runge-kutta 方法来求初始值。

## 变步长的多步方法

- 基本思想：单步方法中用两个精度不一样的方法来估计误差并且通过调整步长来进行误差控制。在多步方法中，显式方法和隐式方法是天然的精度不一样的方法，自然可以采用类似的方法进行误差控制。
- 基本过程：（以采用四阶 Adams-Bashforth technique 和四阶Adams-Moulton technique 方法为例）计算显式方法和隐式方法估计的  $w_{i+1}$  的差，如果大于误差限，则设置新步长为  $qh$  其中

$$q = \left( \frac{270}{19} \frac{h\epsilon}{w_{i+1} - w_{i+1,p}} \right)$$

$w_{i+1,p}$  是 predictor （也就是显式方法）的估计值。同时重新算这个点和前方的三个由 Runge-Kutta 方法计算出来的值。如果前面的点不是通过 Runge-Kutta 方法计算出来的，则不重新计算前面的值，直接用 Runge-Kutta 方法求出新步长下基于前方最后一个点用 Runge-Kutta 算出的三个新值重新预测第四个值。

- 注意点：
  - 隐式方法的估计值是通过在方程右端代入显式方法的估计值得到的。
  - 由于步长的变化，会导致需要重新计算用 Runge-Kutta 方法求得的值的情况。

## 外推法

- 基本思想：与所有的外推法的思路一样，如果误差项形式为

$$R(t, y) = \sum \delta_k h^k$$

其中  $\delta_k$  与步长无关，则可以利用外推法来简单的提高精度。事实上，在一种中点方法

$$w_{i+1} = w_{i-1} + 2hf(t_i, w_i)$$

中进行 "endpoint correction" 后可以做到这一点，误差形式为

$$\sum_{k=1}^{\infty} \delta_k h^{2k}$$

。

- 基本过程：设已知  $w_0 = y(t_0)$ ,  $t_0 = a$ , 想要预测  $y(t_0 + h)$ 。先令步长为  $\frac{h}{2}$  则通过一次中点公式的计算就可以得到  $w_2$

$$w_2 = w_0 + 2h_0 f(a + h_0, w_1)$$

其中中点公式中的  $w_1$  是用欧拉公式计算出来的, 再使用 "endpoint correction", 令

$$y_{1,1} = \frac{1}{2}[w_2 + w_1 + h_0 f(a + 2h_0, w_2)]$$

得到这一步最后的预测。然后调整步长为  $h_1 = \frac{1}{4}$  则有欧拉公式

$$w_1 = w_0 + h f(a, w_0)$$

中点公式

$$w_2 = w_0 + 2h_1 f(a + h_1, w_1)$$

$$w_3 = w_1 + 2h_1 f(a + 2h_1, w_2)$$

$$w_4 = w_2 + 2h_1 f(a + 3h_1, w_3)$$

以及 "endpoint correction"

$$y_{2,1} = w_4 + w_3 + h_1 f(a + 4h_1, w_4)$$

于是我们得到了误差分别为

$$\sum \delta_k \left(\frac{h}{2}\right)^{2k}$$

$$\sum \delta_k \left(\frac{h}{4}\right)^{2k}$$

的预测结果, 将两项按一定比例相减消去误差项中的第一项得到精度更高的估计值. 采用不同的间隔不断重复这个过程, 直到精度达到要求

- 可以选取的步长:  $\{2, 4, 6, 8, 12, 16, 24, 32\}$

## 稳定性

- 一致性和收敛性:
  - 一致性:

$$\lim_{h \rightarrow 0} \max_{1 \leq i \leq N} |\tau_i(h)| = 0$$

其中  $\tau_i(h)$  表示局部截断误差,  $h$  为步长。

- 收敛性:

$$\lim_{h \rightarrow 0} \max_{1 \leq i \leq N} |w_i - y(t_i)| = 0$$

其中  $y(t_i)$  表示真实值。

- 稳定性: 初始值的微小扰动只会带来智慧哥后续的计算带来小的扰动。
- 结论 (单步方法): 对于单步方法

$$w_0 = \alpha$$

$$w_{i+1} = w_i + h\phi(t_i, w_i, h)$$

假设存在  $h_0 > 0$  使得  $\phi(t_i, w_i, h)$  在区间

$$D = \{(t, w, h) | a \leq t \leq b, -\infty \leq w \leq \infty, 0 \leq h \leq h_0\}$$

上连续且对  $w$  满足 Lipschitz 条件, 则

- 该方法稳定
- 差分方法是收敛的当且仅当该方法是连续的, 这等价于

$$\phi(t, y, 0) = f(t, y), \quad \forall a \leq t \leq b$$

- 如果局部截断误差  $\tau_i(h)$  始终小于某一个函数  $\tau(h)$  则

$$|y(t_i) - w_i| \leq \frac{\tau(h)}{L} e^{L(t_i - a)}$$

其中  $L$  是 Lipschitz 条件中的常数。

- 结论（多步方法）：对于多步方法

$$w_0 = \alpha_0, \quad \dots, \quad w_{m-1} = \alpha_{m-1}$$

$$w_{i+1} = a_{m-1}w_i + a_{m-2}w_{i-1} + \dots + a_0w_{i+1-m} + hF(t_i, h, w_{i+1}, w_i, \dots, w_{i+1-m})$$

定义特征多项式是

$$P(\lambda) = \lambda^m - a_{m-1}\lambda^{m-1} - a_{m-2}\lambda^{m-2} - \dots - a_1\lambda - a_0$$

设这个方程的  $m$  个根是  $\lambda_1, \lambda_2, \dots, \lambda_m$  如果  $|\lambda_i| \leq 1, \quad \forall i = 1, 2, \dots, m$  且所有的绝对值为1的根都是单根，就称这个差分方法满足根条件。

- 满足根条件且仅有根中只有唯一的绝对值为 1 的方法称作是强稳定的。
- 满足根条件且仅有根中有多个的绝对值为 1 的方法称作是弱稳定的。
- 不满足根条件的称作是不稳定的。
- 差分方法稳定当且仅当满足根条件；在该方法是 "consistenct" 的时候，差分方法稳定当且仅当该方法收敛。