

## 微分方程数值解——有限差分法

### 色散性和耗散性

参考书目:

- Numerical Partial Differential Equations: Finite Difference Methods (J. W. Thomas, 1995)
- Time Dependent Problems and Difference Methods (B. Gustafsson, 1995)
- Finite Difference Methods for Ordinary and Partial Differential Equations (Randall J. LeVeque, 2007)
- 偏微分方程的有限差分方法 (张强, 2017)

波方程与扩散方程的解具有明显的区别, 这主要表现在两点上:

- 随着时间推进, 扩散方程的解的最大模逐渐减小, 各部分的差距逐渐缩小, 最终趋于平稳 (扩散方程的解可以写成卷积形式); 而波方程的解的最大模保持不变, 各部分的差距保持不变, 解的形状相比初值状态不发生变化。
- 波方程的解可以视作不同单波的组合, 而扩散方程相应的单波解缺乏相应的物理意义。

由于这两点区别的存在, 使得双曲方程的差分格式设计要比抛物方程考虑更多的要求: 为了刻画第一点特性, 即解的形状不发生变化, 需要保证差分格式和问题具有相同的耗散性, 同时希望可以尽可能好地刻画间断点处的变化情况并避免出现数值振荡; 另一方面, 由于第二点的存在, 我们希望差分格式给出的数值波具有和相应的真实单波具有相近的波速、波形以及相位, 还要保持不同单波之间的相对速度和相对位置不发生变化, 这需要差分格式具有与原方程相同的色散性。

本文主要考虑一维常系数对流方程  $u_t + au_x = 0$  的差分格式, 讨论其单调性、耗散性和色散性, 以及如何设计差分格式使得其具有良好的耗散性和色散性。这里处理的差分格式主要为线性显式常系数差分格式, 一般可以表示为

$$u_j^{(n+1)} = \sum_{s=-l}^r \alpha_s u_{j+s}^{(n)}, \quad (0.1)$$

其中  $\alpha_s$  是给定的差分系数。对于一维对流方程  $u_t + au_x = 0$  而言, 直接关于点  $(x_j, t_n)$  做 Taylor 展开可知

$$\Delta t \cdot \tau = (1 - \sum_{s=-l}^r \alpha_s)u - (a\Delta t + \sum_{s=-l}^r s\alpha_s)u_x + (\frac{1}{2}\Delta t^2 a^2 - \frac{1}{2} \sum_{s=-l}^r \alpha_s \Delta x^2 s^2)u_{xx} + O(\Delta t^3) + O(\Delta x^3),$$

于是该格式与原方程相容当且仅当

$$\sum_{s=-l}^r \alpha_s = 1, \quad \sum_{s=-l}^r \alpha_s s = -a \frac{\Delta t}{\Delta x}, \quad (0.2)$$

此时截断误差阶数至少为 1 阶。另外, 当  $a < 0$  时, 该格式具有  $L^2$  稳定性时可以到达的截断误差最高阶数为

$$p = \min(l + r, 2l + 2, 2r), \quad (0.3)$$

并且相应的  $r$  只能为  $l, l+1$  或者  $l+2$  三种情况之一。

## 1 单调性

为了避免出现数值振荡, 差分格式的单调性十分重要。所谓的单调性是指如果某一格点  $u_j^{(n)} \geq u_{j-1}^{(n)}$ , 则下一个时间层内有

$$u_j^{(n+1)} \geq u_{j-1}^{(n+1)}.$$

显然, 单调性是使得数值解形状不发生变化的必要条件, 而对流方程的解  $f(x - at)$  与初值  $f(x)$  仅仅相差一个平移, 因此具有单调保持性质, 自然地我们希望它的数值解也满足单调性。通过考虑单调增加的初值函数

$$f_j = u_j^{(0)} = \begin{cases} 0, & j \leq \ell \\ 1, & j > \ell \end{cases},$$

带入差分格式(0.1)中可得

$$u_1^{(1)} - u_0^{(1)} = \sum_{s=-l}^r \alpha_s [u_{1+s}^{(0)} - u_s^{(0)}] = \alpha_\ell,$$

于是当  $\alpha_\ell \geq 0$ , 则递推可知  $u_1^{(n)} - u_0^{(n)} \geq 0$ , 而当  $\alpha_\ell < 0$  时,  $u_1^{(n)} - u_0^{(n)}$  是正负交替振荡的, 因此我们得到了差分格式单调性的等价定义。

**Definition.** 差分格式(0.1)是单调的, 当且仅当所有的差分系数  $\alpha_\ell \geq 0$ , 因此单调格式也被称作正格式。

差分格式可以保持数值解的单调性, 进而使得数值解不出现数值振荡, 不过单调格式在相容性上有很大的局限性, 即 Godunov 定理。

**Theorem 1. Godunov** 如果差分格式是单调的, 则它至多具有一阶局部截断误差。

*Proof.* 根据之前的分析, 当(0.1)相容时需要满足

$$\sum_{s=-l}^r \alpha_s = 1, \quad \sum_{s=-l}^r \alpha_s s = -a \frac{\Delta t}{\Delta x},$$

此时截断误差的满足

$$\Delta t \cdot \tau = \left( \frac{1}{2} \Delta t^2 a^2 - \frac{1}{2} \sum_{s=-l}^r \alpha_s \Delta x^2 s^2 \right) u_{xx} + O(\Delta t^3) + O(\Delta x^3),$$

根据单调格式的等价定义, 差分系数  $\alpha_s \geq 0$ , 于是根据 Cauchy-Schwarz 不等式有

$$a^2 r^2 = \left( \sum_{s=-l}^r \alpha_s s \right)^2 \leq \left( \sum_{s=-l}^r \alpha_s s^2 \right) \left( \sum_{s=-l}^r \alpha_s \right) = \sum_{s=-l}^r \alpha_s s^2,$$

上式当且仅当  $(-l\sqrt{\alpha_{-l}}, \dots, r\sqrt{\alpha_r})$  与  $(\sqrt{\alpha_{-l}}, \dots, \sqrt{\alpha_r})$  共线时取等号, 此时系数  $\alpha_s$  只能有一个是非零的进而等于 1, 这样的差分格式不具有实际意义, 因此对于有实际意义的差分格式, 上式不取等, 进而  $\tau = O(\Delta t) + O(\Delta x^2/\Delta t)$ , 其局部截断误差至多为一阶。□

这一事实意味着高阶相容的线性格式不可能是单调的, 因此难以避免数值振荡, 为了解决这一问题, 我们需要寻找不具有单调性却又能避免数值振荡的差分格式。

## 1.1 非线性双曲守恒律

考虑一维非线性双曲守恒律:

$$u_t + f(u)_x = 0, \quad (1.1)$$

这类问题的解往往比线性双曲守恒律复杂得多, 可能出现间断解和激波, 为了控制数值解的振荡, 有必要要求差分格式可以保证数值解的单调性。对于这类非线性问题, 我们主要关心它的守恒性格式, 即

**Definition.** 如果某种差分格式形如

$$u_j^{(n+1)} = u_j^{(n)} - \frac{\Delta t}{\Delta x} [F_{j+\frac{1}{2}}^{(n)} - F_{j-\frac{1}{2}}^{(n)}], \quad (1.2)$$

其中

$$F_{j+\frac{1}{2}}^{(n)} = F(u_{j-l+1}^{(n)}, \dots, u_{j+r}^{(n)}) \quad (1.3)$$

称作数值通量, 满足如下要求:

1. Lipschitz 条件 (控制舍入误差的影响):  $F$  关于各个变元都是局部 Lipschitz 连续的;
2. 相容性条件 (保证格式与问题相容):  $F(p, p, \dots, p) = f(p)$ ;

则称该差分格式为守恒性格式。

相比于非守恒形式，守恒型差分格式可以满足局部守恒性质，即

$$\sum_{j=p}^q u_j^{(n+1)} = \sum_{j=p}^q u_j^{(n)} - \frac{\Delta t}{\Delta x} [F_{q+\frac{1}{2}}^{(n)} - F_{p-\frac{1}{2}}^{(n)}], \quad (1.4)$$

进而可以较好的刻画间断解和激波。

最简单的数值通量函数是仅依赖于左右网格点值的两点型数值通量

$$F_{j+\frac{1}{2}}^{(n)} = F(u_j^{(n)}, u_{j+1}^{(n)}).$$

关于守恒型差分格式的数值收敛性，有著名的 Lax-Wendroff 定理：当守恒型差分格式与双曲守恒律相容时，如果网格尺度趋于零时数值解几乎处处有界且收敛到某个函数，则该极限必定是问题的弱解。

## 1.2 单调格式和单调保持格式

对于非线性双曲守恒律，守恒型格式在某种程度上确保了数值解收敛到弱解，但是类似于线性的行波方程，高阶相容和数值振荡的矛盾依然存在。因为双曲守恒律的熵解具有单调保持性质：当初值单调时，熵解具有同样的单调性，所以为了消除数值振荡，我们希望使用的守恒型格式也可以保持解的单调性，为此我们定义所谓的单调保持性质：

**Definition.** 给定单调的初值条件，要求离散得到的数值初值是单调的，如果差分格式给出的任意时刻的数值解都具有与数值初值一样的单调性，则称该差分格式是单调保持格式。

然而这一性质很难验证，注意到相比于单调保持性质，熵解具有更强的比较性质，即考虑两个不同的初值  $u(x, 0)$  和  $v(x, 0)$ ，如果对任意的  $x$  都有  $u(x, 0) \geq v(x, 0)$ ，则对任意的  $t > 0$  都有  $u(x, t) \geq v(x, t)$ ，因此我们希望数值解具有类似的比较性质，即

$$u_j^{(n)} \geq v_j^{(n)} \implies u_j^{(n+1)} \geq v_j^{(n+1)},$$

具有这种性质的格式称为单调格式，事实上，该刻画等价于如下定义：

**Definition.** 设  $l$  和  $r$  是给定的左右臂长，如果数值格式

$$u_j^{(n+1)} = H(u_{j-l}^{(n)}, \dots, u_{j+r}^{(n)}) \quad (1.5)$$

中的函数  $H$  关于每个变元都是单调不减的，则称该差分格式是单调格式。

对于线性格式，单调性和单调保持性是等价的，但是对于非线性格式，单调性和单调保持性是不同的两个概念，一般地，使用如上等价定义可知

$$u_j^{(n)} \geq u_{j-1}^{(n)} \quad \forall j \implies u_j^{(n+1)} \geq u_{j-1}^{(n+1)} \quad \forall j,$$

因此单调格式必然是单调保持格式，但是反之不一定成立。

可以证明单调格式的数值解一致有界，且收敛到双曲守恒律的熵解。然而对于单调格式而言，Godunov 定理依然成立，即相容的单调格式至多具有一阶局部截断误差，因此单调格式不具有高阶相容性。

## 1.3 TVD 格式

由于单调格式存在无法获得高阶相容性的确定，为了提高格式的精度，我们需要做出一些妥协，允许出现轻微的数值振荡，并且希望这种振荡随着网格加密可以逐渐消失。为了量化数值解的振荡强弱，我们引入网格函数的全变差的概念。

**Definition.** 如果  $w : j \mapsto w_j$  是一个网格函数，则称

$$\text{TV}(w) = \sum_j |w_{j+1} - w_j| \quad (1.6)$$

为  $w$  的全变差。

根据 PDE 理论，双曲守恒律的熵解的全变差不随时间增加，因此我们希望差分格式的数值解的全变差也不随时间增加，由此引入了 TVD 格式的概念：

**Definition.** 如果某差分格式的数值解满足

$$\text{TV}(u^{(n+1)}) \leq \text{TV}(u^{(n)}), \quad (1.7)$$

则该差分格式是全变差不增的, 称为 TVD 格式 (Total Variation Diminishing)。

最常用的判断差分格式是 TVD 格式的方法是 Harten 引理:

**Lemma 1. Harten** 设差分格式形如以下增量形式

$$u_j^{(n+1)} = u_j^{(n)} + C_{j+\frac{1}{2}}^+ (u_{j+1}^{(n)} - u_j^{(n)}) - C_{j-\frac{1}{2}}^- (u_j^{(n)} - u_{j-1}^{(n)}), \quad (1.8)$$

如果

$$C_{j+\frac{1}{2}}^+ \geq 0, \quad C_{j-\frac{1}{2}}^- \geq 0, \quad C_{j+\frac{1}{2}}^+ + C_{j-\frac{1}{2}}^- \leq 1, \quad (1.9)$$

则该差分格式是 TVD 格式。

*Proof.* 直接计算可得

$$u_{j+1}^{(n+1)} - u_j^{(n+1)} = (1 - C_{j+\frac{1}{2}}^+ - C_{j+\frac{1}{2}}^-)(u_{j+1}^{(n)} - u_j^{(n)}) + C_{j+\frac{3}{2}}^+(u_{j+2}^{(n)} - u_{j+1}^{(n)}) - C_{j-\frac{1}{2}}^-(u_j^{(n)} - u_{j-1}^{(n)}), \quad (\times)$$

使用三角不等式并重新排列求和指标可知

$$\begin{aligned} \sum_j |u_{j+1}^{(n+1)} - u_j^{(n+1)}| &\leq \sum_j (1 - C_{j+\frac{1}{2}}^+ - C_{j+\frac{1}{2}}^-) |u_{j+1}^{(n)} - u_j^{(n)}| + \sum_j C_{j+\frac{3}{2}}^+ |u_{j+2}^{(n)} - u_{j+1}^{(n)}| + \sum_j C_{j-\frac{1}{2}}^- |u_j^{(n)} - u_{j-1}^{(n)}| \\ &= \sum_j (1 - C_{j+\frac{1}{2}}^+ - C_{j+\frac{1}{2}}^-) |u_{j+1}^{(n)} - u_j^{(n)}| + \sum_j C_{j+\frac{1}{2}}^+ |u_{j+1}^{(n)} - u_j^{(n)}| + \sum_j C_{j-\frac{1}{2}}^- |u_{j+1}^{(n)} - u_j^{(n)}| \\ &= \sum_j |u_{j+1}^{(n)} - u_j^{(n)}|, \end{aligned}$$

所以

$$\text{TV}(u^{(n+1)}) \leq \text{TV}(u^{(n)}),$$

从而该差分格式是 TVD 格式。  $\square$

使用 Harten 引理证明某格式是 TVD 格式的关键在于将数值格式改写为合适的增量形式。

可以证明单调格式都是 TVD 格式, 而 TVD 格式都是单调保持格式, 但是反之不一定成立, TVD 格式不一定是单调格式, 单调保持格式也不一定是 TVD 格式。不过对于线性格式, 单调性、单调保持性质和 TVD 性质是等价的。

Continuous Harten  $\frac{du}{dt} = C_{j+\frac{1}{2}}^+ (u_{j+1} - u_j) - C_{j-\frac{1}{2}}^- (u_j - u_{j-1})$

## 2 耗散性与色散性

现在开始考虑差分格式的耗散性和色散性。如之前所述, 希望差分格式具有和原方程相同的耗散性和色散性, 以便更好地刻画原方程的解, 为此首先需要研究原方程的耗散性和色散性, 之后再讨论差分格式的耗散性和色散性。

$$\Rightarrow \frac{d}{dt} \text{TV}(u) = 0$$

### 2.1 方程的耗散性和色散性

给定某一类型的微分方程, 考虑它如下形式的简谐波解

Fourier in both time and space  $u(x, t) = \hat{u}(\omega, k) e^{i(\omega t + kx)}, \quad (2.1)$

其中  $\omega$  是该谐波的频率,  $k$  是该谐波的波数,  $\lambda = 2\pi/k$  为波长。将谐波解代入原方程, 可以得到

$$\omega = \omega(k), \quad (2.2)$$

该关系称为方程的色散关系或频散关系 (更准确的讲应该是色散-耗散关系, 因为该关系中也包含了方程的耗散信息), 称  $c = -\omega/k$  为该波的波速, 当  $c > 0$  时, 该简谐波从左到右传播, 当  $c < 0$  时, 该简谐波从右到左传播。当考虑多个单波组成的波包时, 该波包的整体波形, 即包络线, 也是一个波函数, 如果考虑一个由波数  $k$  附近的单波组成的波包, 称

$$C_k = -\frac{d\omega(k)}{dk} \quad (2.3)$$

$$u_t = a u_x$$

$$\text{let } u = \hat{u} e^{i\omega t} \Rightarrow \hat{u}_t = a i \omega \hat{u}$$

为该波包的群速度，它表征波包的传播速度，在常规介质内，波包的群速度小于其含有的单波的波速。

现在我们将  $\omega$  的实部和虚部分离，即  $\omega = \omega_r + i\omega_i$ ，此时谐波解为

$$u(x, t) = \hat{u} e^{i(\omega_r t - \omega_i t + kx)} = \hat{u} e^{-\omega_i t} e^{i(\omega_r t + kx)},$$

因此虚部  $\omega_i$  控制了谐波解的衰减速度，因此决定了方程的耗散性，而实部  $\omega_r = \omega_r(k)$  控制了相应波数的谐波解波速，不同波数的谐波解波速间的关系决定了方程的色散性：

### 1. 耗散性：

- 如果存在  $k$  使得  $\omega_i(k) > 0$ ，则谐波解衰减，即方程具有耗散性；
- 如果对所有的波数  $k$  都有  $\omega_i(k) = 0$ ，则谐波解不衰减，即方程不具有耗散性。
- 如果存在  $k$  使得  $\omega_i(k) < 0$ ，则谐波解增长，即方程具有反耗散性。

### 2. 色散性：

- 如果对所有的波数  $k$  而言， $\omega_r(k)$  都是系数保持不变的线性关系，即  $\omega_r(k) = ck$ ，则方程各波数的单波具有相同的波速，整体波形保持不变，因此不具有色散性；
- 如果存在  $k$  使得  $\omega_r(k)$  是非线性关系，或  $c_k = -\omega_r(k)/k$  与  $k$  有关，则方程具有色散性。

**Example.** 考虑一维对流方程  $u_t + au_x = 0$ ，令  $u(x, t) = \hat{u} e^{i(\omega t + kx)}$ ，代入原方程可得

$$i\omega \hat{u} + iak \hat{u} = 0 \implies \omega = -ak,$$

从而得到色散关系  $\omega = -ak$ ，波速为  $c = a$ 。对于热方程  $u_t = \kappa u_{xx}$ ，代入原方程可得

$$i\omega \hat{u} = -\kappa k^2 \hat{u} \implies \omega = i\kappa k^2,$$

因此色散关系变为  $\omega = i\kappa k^2$ 。根据色散关系，可以看出两类方程解的性质有很大的不同：对流方程既不耗散也不色散，因此解的形状保持不变；而热方程同时具有耗散性和色散性。

## 2.2 数值耗散和数值色散

与之前类似，我们考虑差分格式的简谐波解

$$u_j^{(n)} = \hat{u} e^{i(\omega \Delta t n + k \Delta x j)}, \quad (2.4)$$

将其代入差分格式中可得

$$\omega = \omega(k),$$

该关系称作差分格式的离散色散关系，与上一小节类似，令  $\omega = \omega_r + i\omega_i$ ，则差分格式的简谐波解变为

$$u_j^{(n)} = \hat{u} e^{-\omega_i \Delta t n} e^{i(\omega_r \Delta t n + k \Delta x j)},$$

其中  $\omega_i$  和  $\omega_r$  仍然分别决定了差分格式的耗散性和色散性：

### 1. 数值耗散性：

- 如果存在  $k$  使得  $\omega_i(k) > 0$ ，则差分格式具有数值耗散性；
- 如果对所有的波数  $k$  都有  $\omega_i(k) = 0$ ，则差分格式不具有数值耗散性。
- 如果存在  $k$  使得  $\omega_i(k) < 0$ ，则差分格式具有反数值耗散性。

### 2. 数值色散性：

- 如果对所有的波数  $k$  而言， $\omega_r(k)$  都是系数保持不变的线性关系，即  $\omega_r(k) = ck$ ，则差分格式不具有数值色散性；
- 如果存在  $k$  使得  $\omega_r(k)$  是非线性关系，或  $c_k = -\omega_r(k)/k$  与  $k$  有关，则差分格式具有数值色散性。

事实上我们可以利用稳定性分析得到的增长因子来刻画差分格式的耗散性和色散性，这是因为谐波解可以表示为

$$u_j^{(n)} = \hat{u} e^{i(\omega \Delta t n + k \Delta x j)} = \rho^n(k) \hat{u} e^{i k x_j},$$

其中  $\rho(k) = e^{i\omega(k)\Delta t}$  就是差分格式的增长因子，我们有

$$|\rho(k)| = e^{-\omega_i(k)\Delta t}, \quad \arg \rho(k) = \omega_r(k)\Delta t,$$

因此增长因子的模表示  $\Delta t$  之后数值解的振幅衰减情况，而增长因子的幅角表示  $\Delta t$  之后数值解的相位变化情况，并且

$$\omega_i = -\frac{1}{\Delta t} \ln |\rho|, \quad \omega_r = \frac{1}{\Delta t} \arg \rho, \quad (2.5)$$

于是差分格式的增长因子可以给出该格式的耗散性和色散性。

**Example.** 考虑  $[-\pi, \pi]$  上的扩散方程  $u_t = \kappa u_x$  的 FTCS 格式

$$u_j^{(n+1)} = u_j^{(n)} + \frac{\kappa \Delta t}{\Delta x^2} (u_{j+1}^{(n)} - 2u_j^{(n)} + u_{j-1}^{(n)}),$$

代入简谐波解  $u_j^{(n)} = \hat{u}^{(n)} e^{i k x_j}$  进行 Fourier 稳定性分析可得增长因子为

$$\rho(k) = 1 + \frac{\kappa \Delta t}{\Delta x^2} (e^{i k \Delta x} - 2 + e^{-i k \Delta x}) = 1 - 4 \frac{\kappa \Delta t}{\Delta x^2} \sin^2\left(\frac{k \Delta x}{2}\right),$$

其中  $k = 0, 1, \dots, N$ ,  $\Delta x = 2\pi/N$ , 从而得到增长因子的模和幅角分别为

$$|\rho(k)| = \left| 1 - 4 \frac{\kappa \Delta t}{\Delta x^2} \sin^2\left(\frac{k \Delta x}{2}\right) \right|, \quad \arg \rho(k) = 0,$$

因此该差分格式不具有数值色散性，当满足 von Neumann 条件  $\kappa \Delta t / \Delta x^2 \leq 1$  时，该差分格式具有数值耗散性。

一般地，数值色散是数值振荡的根本原因，数值耗散同数值色散的平衡关系决定了数值振荡的具体表现。对于某些问题而言，例如扩散方程，数值耗散是必要的，因此一些格式会人为地引入数值耗散，而另一方面则希望尽可能地避免数值色散，特别当解函数具有尖锐的间断时，数值色散会导致数值解的剧烈振荡。最后，对于多层格式，相应的数值解中一个波数可能对应多个单波，因此数值耗散性和色散性有更复杂的关系，这方面的一个例子可以参见 NPDE-FDM 7.8 节中对蛙跳法的分析。

NPDE. FDM  $\rightarrow$  Modify equation  $\rightarrow$  let  $u = e^{i(\omega t + kx)} \hat{u}$