Problems of Chapter 10.6.1-10.6.4

张志心 计科 2106

日期：2024 年 5 月 6 日

Exercise 10.179

Write down the Butcher tableaux of the modified Euler method, the improved Euler method, and Heun's third-order method in Definition 10.186.

解.

(1) modified Euler method:

$$\begin{array}{c|cc} 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ \hline & 0 & 1 \end{array}$$

(2) improved Euler method:

$$\begin{array}{c|cc} 0 & 0 & 0 \\ 1 & 1 & 0 \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array}$$

(3) Heun's third-order method:

$$\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & 0 & 0 \\ \frac{2}{3} & 0 & \frac{2}{3} & 0 \\ \hline & \frac{1}{4} & 0 & \frac{3}{4} \end{array}$$

□

Exercise 10.180

Verify that the RK method (10.122) can be rewritten as

$$\begin{cases} \boldsymbol{\xi} = \boldsymbol{U}^n + k \sum_{j=1}^{s} a_{i,j} \boldsymbol{f}(\boldsymbol{\xi}_j, t_n + c_j k), \\ \boldsymbol{U}^{n+1} = \boldsymbol{U}^n + k \sum_{j=1}^{s} b_j \boldsymbol{f}(\boldsymbol{\xi}_j, t_n + c_j k), \end{cases} \tag{1}$$

where $i = 1, 2, \ldots, s$.

解. 在 Def 10.177 中令 $\boldsymbol{y}_i = \boldsymbol{f}(\boldsymbol{\xi}_i, t_n + c_i k)$，则

$$\boldsymbol{f}(\boldsymbol{\xi}_i, t_n + c_i k) = \boldsymbol{f}(\boldsymbol{U}^n + k \sum_{j=1}^{s} a_{i,j} \boldsymbol{f}(\boldsymbol{\xi}_j, t_n + c_j k), t_n + c_j k).$$

因此取 $\boldsymbol{\xi}_i = \boldsymbol{U}^n + k \sum_{j=1}^{s} a_{ij} \boldsymbol{f}(\boldsymbol{\xi}_j, t_n + c_j k)$ 则 Definition 10.177 中第一式成立。又由第二式得

$$\boldsymbol{U}^{n+1} = \boldsymbol{U}^n + k \sum_{j=1}^{s} b_j \boldsymbol{f}(\boldsymbol{\xi}_j, t_n + c_j k) = \boldsymbol{U}^n + k \sum_{j=1}^{s} b_j \boldsymbol{y}_j,$$

因此 10.177 第二式也成立。反之，若 Def 10.177 成立，则令 $\boldsymbol{\xi}_i$ 为方程 $\boldsymbol{y}_i = \boldsymbol{f}(\boldsymbol{\xi}_i, t_n + c_i k)$ 的任意一组解，即得本题定义。故本题定义与 10.177 的定义等价。 □

Exercise 10.187

There are three one-parameter families of third-order three-stage ERK methods, one of which is

$$\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \frac{2}{3} & \frac{2}{3} & 0 & 0 \\ \frac{2}{3} & \frac{2}{3} - \frac{1}{4\alpha} & \frac{1}{4\alpha} & 0 \\ \hline & \frac{1}{4} & \frac{3}{4} - \alpha & \alpha \end{array}$$

where α is the free parameter. Derive the above family. Does Heun's third-order formula belong to this family?

解. 设 3-stage ERK 方法的 Butcher tableau 为:

$$\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ c_2 & c_2 & 0 & 0 \\ c_3 & a_{3,1} & a_{3,2} & 0 \\ \hline & b_1 & b_2 & b_3 \end{array}$$

其中 $a_{3,1} + a_{3,2} = c_3$。则

$$\begin{cases} \boldsymbol{y}_1 = \boldsymbol{f}(\boldsymbol{U}^n, t_n), \\ \boldsymbol{y}_2 = \boldsymbol{f}(\boldsymbol{U}^n + k c_2 \boldsymbol{y}_1, t_n + c_2 k), \\ \boldsymbol{y}_3 = \boldsymbol{f}(\boldsymbol{U}^n + k a_{3,1} \boldsymbol{y}_1 + k a_{3,2} \boldsymbol{y}_2, t_n + c_3 k), \\ \boldsymbol{U}^{n+1} = \boldsymbol{U}^n + k(b_1 \boldsymbol{y}_1 + b_2 \boldsymbol{y}_2 + b_3 \boldsymbol{y}_3). \end{cases}$$

由 Example 10.153 有

$$\begin{aligned} \boldsymbol{u}' &= \boldsymbol{f}, \\ \boldsymbol{u}'' &= \boldsymbol{f}_u + \boldsymbol{f}_t, \\ \boldsymbol{u}''' &= \boldsymbol{f}_u^2 \boldsymbol{f} + \boldsymbol{f}_{uu} \boldsymbol{f}^2 + \boldsymbol{f}_u \boldsymbol{f}_t + 2\boldsymbol{f}_{ut} + \boldsymbol{f}_{tt}. \end{aligned}$$

计算截断误差得（以下简记 $\boldsymbol{u}(t_n)=\boldsymbol{u}, \boldsymbol{f}(\boldsymbol{u}(t_n),t_n)=\boldsymbol{f}$，各阶导数和偏导数同理）:

$$
\begin{aligned}
&\mathcal{L}(\boldsymbol{u}(t_n)) \\
=&\boldsymbol{u}(t_{n+1})-\boldsymbol{u}(t_n)-kb_1\boldsymbol{f}(\boldsymbol{u}(t_n),t_n)-kb_2\boldsymbol{f}(\boldsymbol{u}(t_n)+kc_2\boldsymbol{y}_1,t_n+c_2k)-kb_3\boldsymbol{f}(\boldsymbol{u}(t_n)+ka_{3,1}\boldsymbol{y}_1+ka_{3,2}\boldsymbol{y}_2,t_n+c_3k) \\
=&(\boldsymbol{u}+k\boldsymbol{u}'+\frac{k^2}{2}\boldsymbol{u}''+\frac{k^3}{6}\boldsymbol{u}'''+O(k^4))-u-kb_1\boldsymbol{u}'-kb_2(\boldsymbol{f}+kc_2(\boldsymbol{f_u}\boldsymbol{f}+\boldsymbol{f}_t)+\frac{1}{2}k^2c_2^2(\boldsymbol{f}_{tt}+2\boldsymbol{f_{ut}}\boldsymbol{f}+\boldsymbol{f_{uu}}\boldsymbol{f}^2)+O(k^3)) \\
&-kb_3(\boldsymbol{f}+k(a_{3,1}\boldsymbol{f_u}\boldsymbol{f}+a_{3,2}(\boldsymbol{f}+kc_2\boldsymbol{f_u}\boldsymbol{f}+kc_2\boldsymbol{f}_t)\boldsymbol{f_u}+c_3\boldsymbol{f}_t)+\frac{k^2c_3^2}{2}(\boldsymbol{f_{uu}}\boldsymbol{f}^2+2\boldsymbol{f_{ut}}\boldsymbol{f}+\boldsymbol{f}_{tt}))+O(k^4) \\
=&k(1-b_1-b_2-b_3)\boldsymbol{u}'+k^2(\frac{1}{2}-b_2c_2-b_3c_3)\boldsymbol{u}'' \\
&+k^3(\frac{1}{6}\boldsymbol{u}'''-\frac{1}{2}(b_2c_2^2+b_3c_3^2)(\boldsymbol{f_{uu}}\boldsymbol{f}^2+2\boldsymbol{f_{ut}}\boldsymbol{f}+\boldsymbol{f}_{tt})-a_{3,2}b_3c_2(\boldsymbol{f_u}^2\boldsymbol{f}+\boldsymbol{f}_t\boldsymbol{f_u}))+O(k^4)
\end{aligned}
$$

要使该方法达到三阶收敛，需要 $\mathcal{L}(\boldsymbol{u}(t_n))=O(k^4)$。

由 $\boldsymbol{u}$ 的任意性，比较对应系数可知

$$
\begin{aligned}
b_1+b_2+b_3 =&1, \\
b_2c_2+b_3c_3 =&\frac{1}{2}, \\
a_{3,2}b_3c_2 =&\frac{1}{6}, \\
b_2c_2^2+b_3c_3^2 =&\frac{1}{3}.
\end{aligned}
$$

上述方程组有六个未知量但只有四个方程，因此至少有两个自由元。令 $c_3=\frac{2}{3}, b_3=\alpha$，解剩余的方程可得

$$
\begin{aligned}
b_1 &= \frac{1}{4}, \\
b_2 &= \frac{3}{4}-\alpha, \\
c_2 &= \frac{2}{3}, \\
a_{3,2} &= \frac{1}{4\alpha}.
\end{aligned}
$$

即题中给出的 Butcher tableau。Heun's third-stage method 显然不属于这族方法，因为 $c_2=\frac{1}{3}\neq\frac{2}{3}$。 □

Exercise 10.193

Show that the quadrature formula of a RK method is exact for all polynomials f of degree $< r$, i.e.,

$$\forall f\in\mathbb{P}_{r-1}, I_s(f)=\int_{t_n}^{t_n+k} f(t)\mathrm{d}t,$$

if and only if the RK method is $B(r)$.

解. ⇒: 取 $t_n=0, p(t)=t^{l-1}$，因为

$$\frac{k^l}{l}=\int_0^k t^{l-1}\mathrm{d}t=k\sum_{j=1}^s b_j(c_jk)^{l-1}=k^l\sum_{j=1}^s b_jc_j^{l-1}$$

所以

$$\sum_{j=1}^s b_jc_j^{l-1}=\frac{1}{l}.$$

即满足 $B(r)$。

$\Leftarrow$: 只需证明等式对一切 $f(t)=t^l, 0\le l\le r-1$ 成立。即

$$\forall 0\le l\le r-1, k\sum_{j=1}^{s} b_j(t_n+c_jk)^l=\int_{t_n}^{t_n+k} t^l\mathrm{d}t$$

因为 RK method 具有性质 $B(r)$，所以根据 Definition 10.191 有

$$\forall l=1,2,\dots,r,\sum_{j=1}^{s} b_jc_j^{l-1}=\frac{1}{l}.$$

计算可知

$$k\sum_{j=1}^{s} b_j(t_n+c_jk)^l=k\sum_{j=1}^{s} b_j\sum_{m=0}^{l}\binom{l}{m}t_n^m(c_jk)^{l-m}=k\sum_{m=0}^{l}\binom{l}{m}t_n^mk^{l-m}\frac{1}{l-m+1}=\sum_{m=0}^{l}\frac{\binom{l}{m}t_n^mk^{l-m+1}}{l-m+1}$$

另一方面，

$$\int_{t_n}^{t_n+k} t^l\mathrm{d}t=\frac{t^{l+1}}{l+1}\bigg|_{t_n} t_n+k=\frac{1}{l+1}((t_n+k)^{l+1}-t_n^{l+1})=\frac{1}{l+1}\sum_{m=0}^{l}\binom{l+1}{m}t_n^mk^{l-m+1}=\sum_{m=0}^{l}\frac{1}{l-m+1}\binom{l}{m}t_n^mk^{l-m+1}$$

二者相等。故求积式至少有 $r-1$ 阶代数精度。 □

Exercise 10.210

Show that an s-stage collocation method is at least s-order accurate.

解. 只需证明在初值精确的前提下组合方法对任意线性方程 $u'(t)=q(t), q\in\mathbb{P}_{s-1}$ 均精确。$U^0=u(t_0)$。归纳证明 $U^n=u(t_n)$。根据 Definition 10.207 可知，对任意 $i=1,2,\dots,s$，均有

$$p'(t_n+c_ik)=q(t_n+c_ik)=u'(t_n+c_ik).$$

且 $p(t_n)=U^n=u(t_n)$。这是一个关于 u（或 p）的给定 $s+1$ 个条件的 Hermite 插值问题。根据 Definition 10.207, $p\in\mathbb{P}_s$。又因为 $u(t)=U^0+\int_{t_0}^{t} q(t)\mathrm{d}t$, 所以 $u\in\mathbb{P}_s$。s 次插值多项式由 s+1 个插值条件 $u(t_n), q(t_n+c_ik)$ 唯一确定。因此在区间 $[t_n,t_n+k]$ 上 $p=u$。即 $u(t_{n+1})=u(t_n+k)=U^{n+1}=p(t_{n+1})$。因此组合方法对多项式 q 精确。故至少有 s 阶精度。 □

Exercise 10.211

Prove that the collocation method viewed as an RK method satisfies (10.125) and (10.126).

解. 由 Theorem 10.209 可知

$$a_{ij}=\int_0^{c_i} l_j(\tau)\mathrm{d}\tau, b_j=\int_0^1 l_j(\tau)\mathrm{d}\tau.$$

又由 Lemma 2.13（Cauchy relations）可得

$$\sum_{j=1}^{s}\int_0^c l_j(\tau)\mathrm{d}\tau=\int_0^c\sum_{j=1}^{s} l_j(\tau)\mathrm{d}\tau=\int_0^c 1\mathrm{d}\tau=c,\forall c\in\mathbb{R}.$$

所以 $c_i=\sum_{j=1}^{s} a_{ij}, \sum_{j=1}^{s} b_j=1$. □

Exercise 10.213

Derive the three-stage IRK method that corresponds to the collocation points $c_1 = \frac{1}{4}, c_2 = \frac{1}{2}, c_3 = \frac{3}{4}$.

解.

$$
\begin{aligned}
l_1(t) =& \frac{(t-\frac{1}{2})(t-\frac{3}{4})}{(\frac{1}{4}-\frac{1}{2})(\frac{1}{4}-\frac{3}{4})} = 8t^2 - 10t + 3.\\
l_2(t) =& \frac{(t-\frac{1}{4})(t-\frac{3}{4})}{(\frac{1}{2}-\frac{3}{4})(\frac{1}{2}-\frac{3}{4})} = -16t^2 + 16t - 3.\\
l_3(t) =& \frac{(t-\frac{1}{4})(t-\frac{1}{2})}{(\frac{3}{4}-\frac{1}{4})(\frac{3}{4}-\frac{1}{2})} = 8t^2 - 6t + 1.\\
b_1 =& \int_0^1 l_1(\tau)\mathrm{d}\tau = \frac{2}{3}.\\
a_{1,1} =& \int_0^{c_1} l_1(\tau)\mathrm{d}\tau = \frac{23}{48}.\\
a_{2,1} =& \int_0^{c_2} l_1(\tau)\mathrm{d}\tau = \frac{7}{12}.\\
a_{3,1} =& \int_0^{c_3} l_1(\tau)\mathrm{d}\tau = \frac{9}{16}.\\
b_2 =& \int_0^1 l_2(\tau)\mathrm{d}\tau = -\frac{1}{3}.\\
a_{1,2} =& \int_0^{c_1} l_2(\tau)\mathrm{d}\tau = -\frac{1}{3}.\\
a_{2,2} =& \int_0^{c_2} l_2(\tau)\mathrm{d}\tau = -\frac{1}{6}.\\
a_{3,2} =& \int_0^{c_3} l_2(\tau)\mathrm{d}\tau = 0.\\
b_3 =& \int_0^1 l_3(\tau)\mathrm{d}\tau = \frac{2}{3}.\\
a_{1,3} =& \int_0^{c_1} l_3(\tau)\mathrm{d}\tau = \frac{5}{48}.\\
a_{2,3} =& \int_0^{c_2} l_3(\tau)\mathrm{d}\tau = \frac{1}{12}.\\
a_{3,3} =& \int_0^{c_3} l_3(\tau)\mathrm{d}\tau = \frac{3}{16}.
\end{aligned}
$$

因此导出的组合方法的 Butcher tableau 为: □

$$
\begin{array}{c|ccc}
\frac{1}{4} & \frac{23}{48} & -\frac{1}{3} & \frac{5}{48} \\
\frac{1}{2} & \frac{7}{12} & -\frac{1}{3} & \frac{1}{12} \\
\frac{3}{4} & \frac{9}{16} & 0 & \frac{3}{16} \\
\hline
 & \frac{2}{3} & -\frac{1}{3} & \frac{2}{3}
\end{array}
$$

Exercise 10.216

Show $B(s+r)$ and $C(s)$ imply $D(r)$ via multiplying the two vectors $u_j := \sum_{i=1}^{s} b_i c_i^{m-1} a_{i,j}$ and $v_j := \frac{1}{m} b_j (1-c_j^m)$ by the Vandermonde matrix $V(c_1, c_2, \ldots, c_s)$ in Definition 2.3.

解. 设 $(a_{ij}), (b_j), (c_i)$ 满足性质 $B(s+r)$ 和 $C(s)$，令

$$\boldsymbol{u}^{(m)} = [u_1^{(m)}, \ldots, u_s^{(m)}]^T, \boldsymbol{v}^{(m)} = [v_1^{(m)}, \ldots, v_s^{(m)}]^T, u_j^{(m)} = \sum_{i=1}^{s} b_i c_i^{m-1} a_{i,j}, v_j^{(m)} = \frac{1}{m} b_j (1-c_j^m),$$

$$V = V(c_1, c_2, \ldots, c_s) = \begin{bmatrix} 1 & 1 & \ldots & 1 \\ c_1 & c_2 & \ldots c_s & \\ \ldots & \ldots & \ldots & \ldots \\ c_1^{s-1} & c_2^{s-1} & \ldots c_s^{s-1} & \end{bmatrix}.$$

则

$$(V\boldsymbol{u}^{(m)})_j = \sum_{l=1}^{s} c_l^{j-1} \sum_{i=1}^{s} b_i c_i^{m-1} a_{il} = \sum_{i=1}^{s} b_i c_i^{m-1} \sum_{l=1}^{s} c_l^{j-1} a_{il} = \sum_{i=1}^{s} b_i c_i^{m-1} \frac{c_i^j}{j} = \frac{1}{j} \sum_{i=1}^{s} b_i c_i^{m+j-1} = \frac{1}{j(m+j)},$$

且

$$(V\boldsymbol{v}^{(m)})_j = \frac{1}{m} \sum_{l=1}^{s} c_l^{j-1} b_j (1-c_j^m) = \frac{1}{m} (\sum_{l=1}^{s} b_j c_l^{j-1} - \sum_{l=1}^{s} b_j c_l^{m+j-1}) = \frac{1}{m} (\frac{1}{j} - \frac{1}{m+j}) = \frac{1}{j(m+j)}$$

所以对任意的 $m = 1, 2, \ldots, r$，均有 $V\boldsymbol{u}^{(m)} = V\boldsymbol{v}^{(m)}$。因为范德蒙德矩阵 V 一定可逆，所以 $\boldsymbol{u}^{(m)} = \boldsymbol{v}^{(m)}$。根据 Definition 10.214，$D(r)$ 成立。 □

Exercise 10.220

Determine the order of accuracy of the collocation method derived in Exercise 10.213.

解.

$$\begin{aligned} q_r(x) =& (x - \frac{1}{4})(x - \frac{1}{2})(x - \frac{3}{4}) = x^3 - \frac{3}{2}x^2 + \frac{11}{16}x - \frac{3}{32}. \\ & \int_0^1 q_r(x) \mathrm{d}x = 0 \\ & \int_0^1 x q_r(x) \mathrm{d}x = \frac{7}{960} \neq 0 \end{aligned}$$

因此组合求积式的精度为 $3 + 1 = 4$ 阶。 □

参考文献

[1] 张庆海. “Notes on Numerical Analysis and Numerical Methods for Differential Equations”. In: (2024).