第 43 卷第 8 期　　数学的实践与认识　　Vol.43, No.8
2013年 4 月　　MATHEMATICS IN PRACTICE AND THEORY　　Apr., 2013

一组一点外推的数值积分公式的提出与验证

张雨浓, 何良宇, 金 龙, 曲 璐, 刘锦荣

(中山大学 信息科学与技术学院, 广东 广州 510006)

摘 要: 通过利用拉格朗日多项式对等间距的离散采样数据进行函数逼近, 就相等间距外推一点进行积分, 得到了具有一定未来预测能力的一组一点外推的数值积分公式. 计算机数值实验的结果表明了公式的有效性, 即应用公式预测的积分结果可以达到一个较高的精度.

关键词: 预测; 数值积分; 拉格朗日多项式; 一点外推

1 引言

在科学与工程计算如信号处理 [1]、小波变换 [2] 和结构力学 [3] 等领域中, 积分是一个非常重要的计算工具 [4], 但在实际应用中通常无法找到原函数的解析式, 从而不能通过解析法得到精确解, 因此, 采用数值积分得到高精度近似解便成为一种非常重要且常用的方法. 数值积分方法是通过采集指定区间的离散数据点及其相对应的函数值来逼近未知函数的区间积分值. 辛普森公式、龙贝格积分等都是常用的数值积分方法 [5]. 不仅如此, 人们还从积分的精度与收敛速度上进一步研究, 基于神经网络算法 [6] 与人工鱼群算法 [7] 等具有更优性能的新数值积分方法的提出, 显著拓宽了数值积分在实际中的应用. 但是这些积分方法对现有的数据进行处理后只能得到已知区间的函数积分解, 而不能对区间外的积分值加以判定. 随着科学技术的发展, 工程领域已经不再满足于对当下情况的了解, 对未来情况预测能力的不足制约了控制工程领域进一步发展, 于是, 具有前瞻能力的数值积分方法也亟需提出.

针对这类问题, 本文用等间距的已知数据采样点通过拉格朗日多项式进行函数逼近, 以获得已知区间外相等间距外推一点的积分值. 并且, 文中给出了依据 2 至 16 个已知数据点进行一点外推计算的公式, 即一组一点外推的数值积分公式.

2 一组一点外推的数值积分公式

从多项式插值理论出发, 利用已知的区间 $[t_0, t_n]$ 内的一组离散采样数据点及对应的函数值, 构造拉格朗日多项式 $L(t)$ 逼近未知目标函数 $f(t)$, 即 $f(t)\approx L(t)$; 而在实际应用中, 数字离散系统通常是等间距采集数据的 [8], 即 $t_m - t_{m-1} = \Delta t$, $(m = 1,\ 2,\ 3,\ \cdots,\ n)$. 进而, 在外推相同间距一点的区间 $[t_0, t_{n+1}]$ 上对 $L(t)$ 进行积分, 便可近似得到 $f(t)$ 外推一点 t_{n+1} 的数

收稿日期: 2012-11-09
资助项目: 国家自然科学基金 (61075121, 60935001); 国家教育部高等学校博士学科点专项科研基金博导类课题 (20100171110045); 国家大学生创新训练项目 (201210558042)

值积分. 即

$$\int_{t_0}^{t_{n+1}} f(t)\,\mathrm{d}t \approx \int_{t_0}^{t_{n+1}} L(t)\,\mathrm{d}t$$

故本文接下来将主要探讨一组一点外推的数值积分公式.

定义 1[9−11] 设未知目标函数 $f(t)$ 在离散采样数据点 t_i 处的函数值为 $f(t_i)$, $(i = 0,\ 1,\ 2,\ \cdots,\ n)$. 若存在 n 次多项式 $P_n(t) = a_0+a_1t+a_2t^2+\cdots+a_nt^n$ 满足条件 $P_n(t_i) = f(t_i)$, 则称 $P_n(t)$ 为未知目标函数 $f(t)$ 的 n 次插值多项式.

引理 1[9−11] 如果 $t_0,\ t_1,\ t_2,\ \cdots,\ t_n$ 是采样区间 $[a,b]$ 上 $(n+1)$ 个互不相同的数据点, 且未知目标函数 $f(t)$ 的 $(n+1)$ 阶导数连续 (记 $f \in C^{(n+1)}[a,b]$), 则对任一 $t \in [a,b]$, 存在 $c \in (a,b)$ 使得 $f(t) = P_n(t) + R_{n+1}(t)$, 其中 $P_n(t)$ 为未知目标函数 $f(t)$ 的插值多项式, $R_{n+1}(t)$ 为余项且 $R_{n+1}(t) := f^{(n+1)}(c)(t-t_0)(t-t_1)\cdots(t-t_n)/(n+1)!$.

根据定义 1 与引理 1, 令 $P_n(t) = L(t)$, 可得

$$f(t) = L(t) + R_{n+1}(t)$$

等式两边分别在区间 $[t_0, t_{n+1}]$ 内进行积分

$$\int_{t_0}^{t_{n+1}} f(t)\,\mathrm{d}t = \int_{t_0}^{t_{n+1}} L(t)\,\mathrm{d}t + \int_{t_0}^{t_{n+1}} R_{n+1}(t)\mathrm{d}t$$

根据拉格朗日多项式逼近定理 [5]

$$\int_{t_0}^{t_{n+1}} L(t)\mathrm{d}t = \int_{t_0}^{t_{n+1}} \sum_{k=0}^{n} f(t_k)\frac{\prod_{\substack{j=0\\ j\neq k}}^{n}(t-t_j)}{\prod_{\substack{j=0\\ j\neq k}}^{n}(t_k-t_j)}\mathrm{d}t \tag{1}$$

令 $t = t_0 + \Delta t \times x$, 其中 $x \geq 0$; 令 $t_l = t_0 + \Delta t \times l$, 其中 $l = k$ 和 j; 则有

$$\begin{aligned}&\int_{t_0}^{t_{n+1}} \sum_{k=0}^{n} f(t_k)\frac{\prod_{\substack{j=0\\ j\neq k}}^{n}(t-t_j)}{\prod_{\substack{j=0\\ j\neq k}}^{n}(t_k-t_j)}\mathrm{d}t\\ &=\sum_{k=0}^{n} f(t_k)\int_{0}^{n+1}\frac{\prod_{\substack{j=0\\ j\neq k}}^{n}[(t_0+\Delta t\times x)-(t_0+\Delta t\times j)]}{\prod_{\substack{j=0\\ j\neq k}}^{n}[(t_0+\Delta t\times k)-(t_0+\Delta t\times j)]}\,d(t_0+\Delta t\times x)\end{aligned}$$

忽略余项, 化简得到一组一点外推的数值积分公式

$$\int_{t_0}^{t_{n+1}} f(t)\mathrm{d}t \approx \sum_{k=0}^{n}\Delta t\times f(t_k)\int_{0}^{n+1}\frac{\prod_{\substack{j=0\\ j\neq k}}^{n}(x-j)}{\prod_{\substack{j=0\\ j\neq k}}^{n}(k-j)}\,dx \tag{2}$$

则, 包括待积分区间两个边界点在内, 等间距地选取待积分区间内的 $N+1$ 个数据点 (其中 $N = n+1$), 利用前 N 个点作为一组已知数据点, 使用公式 (2) 可以外推一点对整个区间进行积分.

分别令 $n = 1,\ 2,\ \cdots,\ 15$, 经过计算, 即可得到 $2 \sim 16$ 个一组等间距已知数据点的一组一点外推的数值积分公式. 以下列出 $2 \sim 16$ 个数据点的一组一点外推的数值积分公式, 以方便其他研究者和应用者的尝试与探讨. [注: 用 f_n 代替 $f(t_n)$ 以简写公式]

$$\int_{t_0}^{t_2} f(t)\mathrm{d}t \approx 2\Delta t f_1$$

$$\int_{t_0}^{t_3} f(t)\mathrm{d}t \approx \Delta t\left(\frac{3}{4}f_0 + \frac{9}{4}f_2\right)$$

$$\int_{t_0}^{t_4} f(t)\mathrm{d}t \approx \Delta t\left(\frac{8}{3}f_1-\frac{4}{3}f_2+\frac{8}{3}f_3\right)$$

$$\int_{t_0}^{t_5} f(t)\mathrm{d}t \approx \Delta t\left(\frac{95}{144}f_0-\frac{25}{72}f_1+\frac{25}{6}f_2-\frac{175}{72}f_3+\frac{425}{144}f_4\right)$$

$$\int_{t_0}^{t_6} f(t)\mathrm{d}t \approx \Delta t\left(\frac{33}{10}f_1-\frac{21}{5}f_2+\frac{39}{5}f_3-\frac{21}{5}f_4+\frac{33}{10}f_5\right)$$

$$\int_{t_0}^{t_7} f(t)\mathrm{d}t \approx \Delta t\left(\frac{965}{1586}f_0-\frac{49}{72}f_1+\frac{1011}{146}f_2-\frac{1274}{135}f_3+\frac{8811}{743}f_4-\frac{2107}{360}f_5+\frac{1571}{439}f_6\right)$$

$$\int_{t_0}^{t_8} f(t)\mathrm{d}t \approx \Delta t\left(\frac{736}{189}f_1-\frac{848}{105}f_2+\frac{1952}{105}f_3-\frac{10575}{508}f_4+\frac{1952}{105}f_5-\frac{848}{105}f_6+\frac{736}{189}f_7\right)$$

$$\int_{t_0}^{t_9} f(t)\mathrm{d}t \approx \Delta t\left(\frac{260}{453}f_0-\frac{1213}{1211}f_1+\frac{3111}{298}f_2-\frac{4898}{221}f_3+\frac{10287}{280}f_4-\frac{4305}{121}f_5+\frac{3725}{143}f_6\right.$$
$$\left.-\frac{2801}{274}f_7+\frac{2515}{604}f_8\right)$$

$$\int_{t_0}^{t_{10}} f(t)\mathrm{d}t \approx \Delta t\left(\frac{2109}{473}f_1-\frac{4175}{324}f_2+\frac{20911}{569}f_3-\frac{5524}{91}f_4+\frac{9644}{129}f_5-\frac{5524}{91}f_6+\frac{20911}{569}f_7\right.$$
$$\left.-\frac{4175}{324}f_8+\frac{2109}{473}f_9\right)$$

$$\int_{t_0}^{t_{11}} f(t)\mathrm{d}t \approx \Delta t\left(\frac{859}{1566}f_0-\frac{4705}{3582}f_1+\frac{5929}{404}f_2-\frac{3113}{74}f_3+\frac{10269}{115}f_4-\frac{8109}{65}f_5+\frac{46835}{364}f_6\right.$$
$$\left.-\frac{9814}{107}f_7+\frac{35555}{734}f_8-\frac{4741}{306}f_9+\frac{4574}{969}f_{10}\right)$$

$$\int_{t_0}^{t_{12}} f(t)\mathrm{d}t \approx \Delta t\left(\frac{9626}{1925}f_1-\frac{4627}{249}f_2+\frac{7799}{122}f_3-\frac{13972}{101}f_4+\frac{12005}{54}f_5-\frac{7956}{31}f_6+\frac{12005}{54}f_7\right.$$
$$\left.-\frac{13972}{101}f_8+\frac{7799}{122}f_9-\frac{4627}{249}f_{10}+\frac{9626}{1925}f_{11}\right)$$

$$\int_{t_0}^{t_{13}} f(t)\mathrm{d}t \approx \Delta t\left(\frac{921}{1742}f_0-\frac{8688}{5369}f_1+\frac{5569}{284}f_2-\frac{6280}{89}f_3+\frac{1657}{9}f_4-\frac{19679}{59}f_5+\frac{21255}{47}f_6\right.$$
$$\left.-\frac{22751}{50}f_7+\frac{16998}{49}f_8-\frac{8726}{45}f_9+\frac{1371}{17}f_{10}-\frac{26691}{1234}f_{11}+\frac{1855}{353}f_{12}\right)$$

$$\int_{t_0}^{t_{14}} f(t)\mathrm{d}t \approx \Delta t\left(\frac{1497}{271}f_1-\frac{4750}{189}f_2+\frac{67834}{667}f_3-\frac{16005}{59}f_4+\frac{14061}{26}f_5-\frac{52339}{65}f_6+\frac{21187}{23}f_7\right.$$
$$\left.-\frac{52339}{65}f_8+\frac{14061}{26}f_9-\frac{16005}{59}f_{10}+\frac{67834}{667}f_{11}-\frac{4750}{189}f_{12}+\frac{1497}{271}f_{13}\right)$$

$$\int_{t_0}^{t_{15}} f(t)\mathrm{d}t \approx \Delta t\left(\frac{1239}{2417}f_0-\frac{439}{229}f_1+\frac{4839}{194}f_2-\frac{7195}{66}f_3+\frac{28097}{83}f_4-\frac{21803}{29}f_5+\frac{16488}{13}f_6\right.$$
$$-\frac{37765}{23}f_7+\frac{137510}{83}f_8-\frac{25947}{20}f_9+\frac{29140}{37}f_{10}-\frac{10475}{29}f_{11}+\frac{6584}{53}f_{12}$$
$$\left.-\frac{5120}{179}f_{13}+\frac{2205}{382}f_{14}\right)$$

$$\int_{t_0}^{t_{16}} f(t)\mathrm{d}t \approx \Delta t\left(\frac{1490}{247}f_1-\frac{3381}{104}f_2+\frac{20315}{134}f_3-\frac{25921}{54}f_4+\frac{42351}{37}f_5-\frac{77127}{37}f_6+\frac{44581}{15}f_7\right.$$

$$-\frac{100157}{30}f_8+\frac{44581}{15}f_9-\frac{77127}{37}f_{10}+\frac{42351}{37}f_{11}-\frac{25921}{54}f_{12}+\frac{20315}{134}f_{13}$$
$$-\frac{3381}{104}f_{14}+\frac{1490}{247}f_{15}\Big)$$

3 计算机数值实验

为了验证一组一点外推的数值积分公式的计算效果, 本文选取了目标函数 $y=x^2\mathrm{e}^x$ 与 $y=\sin x$ 进行计算机数值实验. 考虑到积分区间和外推距离的限制 [5] , 分别选取了 $[0,0.02]$, $[0,0.05]$, $[0,0.1]$, $[0,0.2]$, $[0,0.5]$ 五个积分区间进行 $2\sim16$ 个点的数值验证, 比较分析估计值与真实值的相对误差, 结果分别如表 1 与表 2 所示. [注: 粗体数据表示每个积分区间内能够取得的最小误差]

表 1 目标函数 $y=x^2\mathrm{e}^x$ 在 5 个积分区间内利用 $2\sim16$ 个数据点进行一组一点外推积分的计算误差

	[0,0.02]	[0,0.05]	[0,0.1]	[0,0.2]	[0,0.5]
2	2.53×10^{-1}	2.59×10^{-1}	2.69×10^{-1}	2.87×10^{-1}	3.41×10^{-1}
3	1.67×10^{-3}	4.20×10^{-3}	8.48×10^{-3}	1.72×10^{-2}	4.52×10^{-2}
4	4.38×10^{-6}	2.75×10^{-5}	1.10×10^{-4}	4.44×10^{-4}	2.83×10^{-3}
5	1.02×10^{-8}	1.58×10^{-7}	1.26×10^{-6}	1.01×10^{-5}	1.56×10^{-4}
6	$\mathbf{4.40\times10^{-12}}$	5.92×10^{-10}	9.37×10^{-9}	1.49×10^{-7}	5.71×10^{-6}
7	1.06×10^{-11}	7.57×10^{-12}	6.71×10^{-11}	2.09×10^{-9}	1.98×10^{-7}
8	1.07×10^{-11}	5.51×10^{-12}	1.55×10^{-12}	2.20×10^{-11}	5.15×10^{-9}
9	1.06×10^{-11}	5.51×10^{-12}	1.21×10^{-12}	3.27×10^{-13}	1.30×10^{-10}
10	1.07×10^{-11}	5.51×10^{-12}	$\mathbf{1.20\times10^{-12}}$	1.10×10^{-13}	2.64×10^{-12}
11	1.07×10^{-11}	5.51×10^{-12}	1.22×10^{-12}	1.02×10^{-13}	5.80×10^{-14}
12	1.06×10^{-11}	5.51×10^{-12}	1.22×10^{-12}	1.09×10^{-13}	1.50×10^{-14}
13	1.06×10^{-11}	$\mathbf{5.49\times10^{-12}}$	1.22×10^{-12}	1.01×10^{-13}	$\mathbf{3.00\times10^{-15}}$
14	1.07×10^{-11}	5.51×10^{-12}	1.21×10^{-12}	1.06×10^{-13}	1.20×10^{-14}
15	1.06×10^{-11}	5.53×10^{-12}	1.23×10^{-12}	$\mathbf{8.10\times10^{-14}}$	3.30×10^{-14}
16	1.07×10^{-11}	5.54×10^{-12}	1.25×10^{-12}	1.75×10^{-13}	5.80×10^{-14}

从表 1 与表 2 可以看出, 一组一点外推数值积分公式所得到的结果与标准结果相比, 通过选取合适的数据点个数和积分区间, 可以实现对目标函数的高精度估计 (如相对误差可达到 $10^{-11}\sim10^{-15}$ 的数量级). 从目标函数 $y=x^2\mathrm{e}^x$ 的计算机数值实验结果可以发现: 在 $[0,0.02]$ 区间中选用数据点个数为 6 时, 误差精度就已经达到 10^{-12} 的数量级; 而选用更多的数据点则使目标函数的误差波动范围变小, 误差精度趋于稳定. 而积分区间 $[0,0.05]$, $[0,0.1]$, $[0,0.2]$, $[0,0.5]$ 分别相应地在 7, 8, 9, 11 个数据点的时候达到了误差精度相对稳定的状态. 一般而言, 采用的数据点越多, 计算机数值实验得到的结果精度越高, 但是当精度达到一定的数量级之后, 数据点的增多会造成误差在一个较小的区间内波动, 并不能有效地减小误差. 从目标函数 $y=x^2\mathrm{e}^x$ 的计算机数值实验结果还可以发现: 积分区间 $[0,0.02]$ 内的一组一点外推数值积分误差精度可达到 10^{-11} 的数量级, 而 $[0,0.05]$, $[0,0.1]$, $[0,0.2]$, $[0,0.5]$ 四个积分区间的

误差精度可分别达到 10^{-12}, 10^{-12}, 10^{-13}, 10^{-14} 的数量级. 从中可以得出, 随着区间长度的增加, 达到相对稳定的误差精度所需要的数据点个数会相应增加, 而稳定时的误差精度也有提高的趋势.

表 2 目标函数 $y=\sin x$ 在 5 个积分区间内利用 2～16 个数据点进行一组一点外推积分的计算误差

	[0,0.02]	[0,0.05]	[0,0.1]	[0,0.2]	[0,0.5]
2	1.67×10^{-5}	1.04×10^{-4}	4.17×10^{-4}	1.67×10^{-3}	1.05×10^{-2}
3	3.70×10^{-6}	2.31×10^{-5}	9.26×10^{-5}	3.70×10^{-4}	2.29×10^{-3}
4	4.87×10^{-11}	1.90×10^{-9}	3.04×10^{-8}	4.87×10^{-7}	1.91×10^{-5}
5	6.87×10^{-12}	2.64×10^{-10}	4.22×10^{-9}	6.74×10^{-8}	2.61×10^{-6}
6	1.11×10^{-13}	4.50×10^{-14}	1.04×10^{-12}	6.70×10^{-11}	1.65×10^{-8}
7	1.10×10^{-13}	3.10×10^{-14}	9.50×10^{-14}	6.74×10^{-12}	1.63×10^{-9}
8	1.10×10^{-13}	2.90×10^{-14}	1.10×10^{-14}	5.00×10^{-15}	8.20×10^{-12}
9	1.11×10^{-13}	2.90×10^{-14}	1.20×10^{-14}	$\mathbf{1.00\times10^{-15}}$	6.37×10^{-13}
10	1.11×10^{-13}	2.90×10^{-14}	1.20×10^{-14}	3.00×10^{-15}	6.00×10^{-15}
11	1.13×10^{-13}	2.80×10^{-14}	$\mathbf{1.00\times10^{-14}}$	1.00×10^{-15}	**0**
12	1.17×10^{-13}	$\mathbf{1.80\times10^{-14}}$	2.00×10^{-14}	3.00×10^{-15}	7.00×10^{-15}
13	1.21×10^{-13}	2.70×10^{-14}	1.60×10^{-14}	9.00×10^{-15}	4.00×10^{-15}
14	1.25×10^{-13}	3.20×10^{-14}	2.50×10^{-14}	1.60×10^{-14}	1.40×10^{-14}
15	$\mathbf{9.40\times10^{-14}}$	4.00×10^{-14}	4.10×10^{-14}	6.00×10^{-15}	8.00×10^{-15}
16	1.17×10^{-13}	5.70×10^{-14}	3.50×10^{-14}	6.20×10^{-14}	6.50×10^{-14}

从目标函数 $y=\sin x$ 的计算机验证结果中, 也可以发现在积分区间 $[0,0.02]$ 内选取 6 个数据点的时候, 误差精度就可以稳定在 10^{-13} 的数量级, 而 $[0,0.05]$, $[0,0.1]$, $[0,0.2]$, $[0,0.5]$ 四个积分区间分别在选用 6, 7, 8, 9 个数据点的时候达到相对稳定的误差精度, 分别是 10^{-14}, 10^{-14}, 10^{-15}, 10^{-15} 的数量级. 计算机数值实验结果表明了一组一点外推的数值积分公式的有效性, 即在预测的结果中, 误差可以达到一个较高的精度.

4 结论

数值积分在科学技术与工程应用的领域中都发挥着重要作用, 但是现有的常用数值积分公式缺乏对已知区间外的积分计算能力, 这限制了数值积分在工程实践上的进一步应用. 一组一点外推的数值积分公式便是在这个背景下提出来的. 本文利用等间距离散数据点得到拉格朗日多项式来逼近未知函数, 进而进行带预测性质的积分. 计算机数值实验结果表明, 这种新的数值积分方法具有有效的预测性能和优良的精度.

参考文献

[1] 陈为真, 汪秉文, 胡晓娅. 基于时域积分的加速度信号处理 [J]. 华中科技大学学报, 2010, 38(1): 1-4.

[2] 梁锡坤. 连续小波变换的 Thiele 型算法 [J]. 数值计算与计算机应用, 2004(2): 116-121.

[3] 张森文, 曹开彬. 计算结构动力响应的状态方程直接积分法 [J]. 计算机力学学报, 2000, 17(1): 94-97.

[4] 刘彩艳, 陈兴同. 数值积分的双侧逼近及应用 [J]. 大学数学, 2012, 28(1): 61-66.

[5] Mathews J H, Fink K D. Numerical Methods Using MATLAB[M]. 北京: 电子工业出版社, 2005.

[6] 罗玉雄, 文卉. 一种基于神经网络算法的数值积分方法 [J]. 传感技术学报, 2006, 19(4): 1187-1189.

[7] 聂黎明, 周永权. 基于人工鱼群算法求任意函数的数值积分 [J]. 数学的实践与认识, 2009, 39(19): 127-134.

[8] Franklin G F, Powell J D, Naeini A E. 自动控制原理与设计 [M]. 李中华, 张雨浓译. 第 5 版. 北京: 人民邮电出版社, 2007.

[9] 张雨浓, 郭东生, 徐思洪, 等. 未知函数之一阶数值微分公式验证与实践 [J]. 甘肃科学学报, 2009, 21(1): 13-18.

[10] 张雨浓, 陈宇曦, 陈锦浩, 等. 一阶超前数值差分公式的提出, 研究与实践 [J]. 中山大学学报, 2010, 51(2): 1-5.

[11] 李建良, 蒋勇, 汪光先. 计算机数值方法 [M]. 南京: 东南大学出版社, 2000.

Proposing and Verification of a Group of 1-Node-Extrapolated Numerical Integration Formulas

ZHANG Yu-nong, HE Liang-yu, JIN Long, QU Lu, LIU Jin-rong

(School of Information Science and Technology, Sun Yat-sen University, Guangzhou 510006, China)

Abstract: In this paper, the Lagrange polynomial is used to approximate the unknown target function based on equally-based discrete-time sampling data. In order to calculate the integration of the interval including an extrapolated node, a group of 1-node-extrapolated numerical integration formulas are proposed, which are capable of predicting the future information. Numerical results show the efficacy of the formulas, that is, the prediction results show high accuracies.

Keywords: prediction; integration; Lagrange polynomial; 1-node-extrapolated