

游性恬

(北京气象学院, 100081)

提 要

该文首先引进 3 种精度较高的边界条件, 借助于 Spline 函数, 近似计算离散网格上的一阶空间微商, 并通过数值试验进行比较, 然后用 Lagrangian、Spline 和 Hermite 方法对不同相对位置的插值点作插值计算, 并给出了一套较为简单适用的 Hermite 方案计算公式。结果表明, 这 3 类插值方法都能达到一定的精度, 但其中以取周期性边界条件, 用 Hermite 方法算得的插值函数最为精确, 其误差约为采用二阶或三阶单向差边条件相应误差的 1/4 或 1/2, 比 Lagrangian 方法的相应值要小一个量级。

关键词: 插值 边界条件 误差

前 言

数值插值计算在数值天气预报、尤其是套网格模式的实现、运行中有重要的作用。插值的精度直接影响数值预报的准确性^[1-3]。在物理量诊断场的计算中也经常要用到数值插值方法。由于在 Spline 和 Hermite 插值方法中需要用到插值函数在网格点上的一阶空间微商值, 并且一般是在人为给定的边界条件下通过 Spline 函数计算之, 而以往采用精度较低的一阶微商来外推边界条件^[4]对空间微商的计算往往带来较大的误差, 因此我们提出改用 3 种更为精确的边界条件, 并进行了具体计算和比较; 又由于数值插值的误差还与插值点相对于网格点的位置有关, 因此, 我们以正弦函数为例, 系统地分析比较了 3 种常用的插值方案^[5], 即 Lagrangian (拉格朗日)、Spline (样条) 和 Hermite (埃尔米特) 插值方案 (以下分别简称为 L、S、H 方案), 分析它们对于不同相对位置的插值点以及取不同近似的空间微商情况下插值函数的误差特点, 并对其中的 H 方案给出一套较为简单、

适用的计算公式。

1 准确解

为简单起见, 本文用作数值试验的准确解取为一维正弦波动函数 $F(x)$

$$F(x) = A \sin \frac{6\pi x}{L} \quad (1)$$

它在网格点上的值为

$$F_i = A \sin \frac{6\pi x_i}{L} \quad (2)$$

这里设 F_i 代表 500hPa 等压面高度距平值, 其最大振幅 A 取作 400m; 波长 L 取作 6000km; $x_i = id$, 网格距 $d = 200$ km, 网格点 $i = 0, 1, 2, \dots, 30$ 。这样 x 最大值为 6000km, 为整个波长范围。

根据插值点相对于网格点的不同位置, 插值点 p_i 分别取为以下 4 种情况及相应的 i 值 (见图 1):

$$\begin{aligned} A: p_i &= x_i + 0.25d, i = 0, 1, \dots, 29 \\ B: p_i &= x_i + 0.5d, i = 0, 1, \dots, 29 \\ C: p_i &= x_i + 0.75d, i = 0, 1, \dots, 29 \\ D: p_i &= x_i - 0.25d, i = 1, 2, \dots, 29 \end{aligned} \quad (3)$$

1) 本项工作得到廖洞贤先生的指导和八五暴雨攻关课题 85-906-04-02 的部分资助。

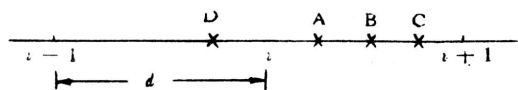


图1 四类插值点的位置

2 一阶空间微商的计算

由于在 S 方案和 H 方案的插值公式中都包含有一阶空间微商值,因此我们首先利用 Spline 函数来计算它。令

$$m_i = \left(\frac{dF}{dx}\right)_i \quad (4)$$

根据一维三次 Spline 函数的光滑性条件有^[5]

$$m_{i-1} + 4m_i + m_{i+1} = \frac{3}{d}(F_{i+1} - F_{i-1}), \quad i = 1, \dots, 29 \quad (5)$$

(3) 周期性边界条件^[7], 即设 $F_0 = F_{30}, m_0 = m_{30}, m'_0 = m'_{30}$

将以上 3 组边界条件分别代入方程组 (5) 中, 可以求得相应的 3 组 m_i 近似值, 依次记为 $m1_i, m2_i, m3_i (i=0, 1, \dots, 30)$ 。以下若无说明, i 取值均同此。对式 (1) 作微商还可算出 m 在网格点上的准确值 ma_i , 进而计算 $m1_i, m2_i$ 和 $m3_i$ 的误差 (误差 = 近似值 - 准确值, 下同), 分别记为 $Em1_i, Em2_i$ 和 $Em3_i$ 。

表1 m 的准确值 ma 及误差 $Em1, Em2, Em3$ 的统计特征量

	AAV	RMS	MAX	MIN
ma	0.828E-3	0.903E-3	0.126E-2	-0.126E-2
$Em1$	0.132E-4	0.379E-4	0.143E-3	-0.397E-4
$Em2$	0.509E-5	0.135E-4	0.509E-4	-0.149E-4
$Em3$	0.750E-6	0.818E-6	0.114E-5	-0.114E-5

图 2 为 $Em1, Em2, Em3$ 在左边界附近的变化 (在右边界为对称分布, 图略)。由图 2 看到, 大的误差出现在两端边界上。由于人为给定的边界值与内点上的计算值不协调, 导

该方程组加上适当的边界条件以后, 即为对角占优的三对角方程组, 可以用追赶法迭代求出 $m_1, \dots, m_{29}$ 。但是人为给定的边界条件, 由于与由模式计算的点值不协调, 往往在边界上形成高频反射波并向里传播, 导致大的误差。常用的几种水平侧边界条件^[6]可能引起预报范围缩小、或歪曲波动在边界附近的状况甚至使之变形等问题; 同时, 由一阶精度的单向差公式外推边界值方法^[4]精度偏低。为此我们改用以下 3 种较高精度的边界条件。

(1) 由二阶精度单向差公式可导出

$$\begin{cases} m_0 = \frac{1}{2d}(-3F_0 + F_1 - F_2) \\ m_{30} = \frac{1}{2d}(F_{28} - 4F_{29} + 3F_{30}) \end{cases} \quad (6)$$

(2) 由三阶精度单向差公式可导出

$$\begin{cases} m_0 = \frac{1}{6d}(-11F_0 + 18F_1 - 9F_2 + 2F_3) \\ m_{30} = \frac{1}{6d}(-2F_{27} + 9F_{28} - 18F_{29} + 11F_{30}) \end{cases} \quad (7)$$

表 1 给出了一阶空间微商 m 的准确值及误差的某些统计特征量, 包括平均绝对值 AAV, 均方根 RMS, 最大值 MAX 和最小值 MIN。从表 1 看到, 近似值 $m1, m2, m3$ 都达到一定的精度, 其中误差最大的 $m1$, 其 AAV 和 RMS 的相对误差 (即 $Em1/ma$) 也在 5% 以下; 而误差最少的 $m3$, 其各项统计特征量的相对误差都在 1‰ 以下, 即 $Em3$ 比 ma 小 3 个量级。

致 $Em1$ 和 $Em2$ 在边界附近产生明显的高频振荡。 $Em3$ 由于采用周期性边界条件, 且其周期与准确解三角函数的固有周期一致, 因而在边界附近变化平缓, 误差最小。

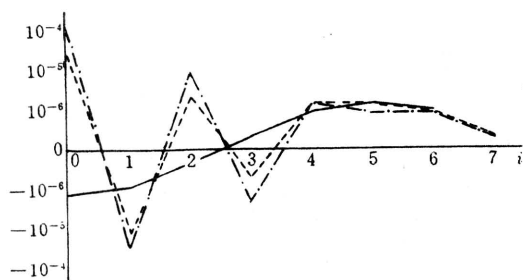


图2 在左边界附近 m 的误差分布
实线: Em_3 断线: Em_2 点划线: Em_1

3 插值函数的计算方案

首先将以上 A、B、C、D 四类插值点(见图1)分别代入式(2),算出相应的插值函数的准确值,记为 FA_i 。然后按以下三种方案,分别计算各类插值点上函数的近似值及其误差。

3.1 L 方案

对于每个插值点,取其附近的4个网格点上的函数值构造三次 Lagrangian 多项式,从而求出插值函数的近似值 Fl_i 及其误差 El_i (对各类插值点, i 的取值见式(3)。以下 S、H 方案同。)

3.2 S 插值

在每个小区间 $[x_i, x_{i+1}]$ 上构造三次样条函数,这是一个三次多项式,且在小区间的两端点上函数值和一阶微商值都等于已知值。因此三次样条函数的光滑性好;且相对于 L 方案而言,为求每个插值函数,只用到较少的格点函数值和微商值^[5](具体公式略)。由于在 S 方案中要用到一阶空间微商 m ,我们将前面算出的 $m1_i, m2_i, m3_i$ 分别代入,得到相应的插值函数值 $Fs1_i, Fs2_i, Fs3_i$ 及其误差 $Es1_i, Es2_i, Es3_i$ 。

3.3 H 插值

H 插值和 S 插值基本思路完全相同。我们根据 H 插值的一般原理和公式,改写成以下简便、适用的形式。

设归一化的插值点为

$$H_i = (p_i - x_i)/d \quad (8)$$

其中 p_i 和 x_i 的定义见式(3),相应的三次 H 插值函数为

$$FH_i = (a_0)_i + (a_1)_i H_i + (a_2)_i H_i^2 + (a_3)_i H_i^3 \quad (9)$$

系数 $(a_0)_i, (a_1)_i, (a_2)_i, (a_3)_i$ 由以下各式确定:

当 $x_i \leq x \leq x_{i+1}$ 时

$$\begin{cases} (a_0)_i = F_i \\ (a_1)_i = dm_i \\ (a_2)_i = 3[F_{i+1} - (a_0)_i] - 2(a_1)_i - dm_{i+1} \\ (a_3)_i = F_{i+1} - (a_0)_i - (a_1)_i - (a_2)_i \end{cases} \quad (10)$$

当 $x_{i-1} \leq x \leq x_i$ 时

$$\begin{cases} (a_0)_i = F_i \\ (a_1)_i = dm_i \\ (a_2)_i = 3[F_{i-1} - (a_0)_i] + 2(a_1)_i + dm_{i-1} \\ (a_3)_i = -F_{i-1} + (a_0)_i - (a_1)_i + (a_2)_i \end{cases} \quad (11)$$

将前面算出的 $m1_i, m2_i, m3_i$ 和 ma_i 分别代入式(10)和(11)中,再利用式(9)求出相应的 H 插值函数值 $FH1_i, FH2_i, FH3_i$ 和 FHA_i 及其误差 $EH1_i, EH2_i, EH3_i$ 和 EHA_i 。显然, EHA_i 反映该方案本身固有的误差,与计算空间微商时所取的边界条件无关。

4 分析和讨论

由于 S 和 H 插值的计算都是应用 Hermite 基函数来构造插值多项式,且两者在各小区间边界上的连接条件也取作相同,因此由这两种方法算出的结果非常接近。两者各统计特征量相差都在 5% (绝对值相差在 $10^{-4}m$) 以下。因此,下面我们只讨论 L 和 H 插值的计算结果。

4.1 计算误差的比较

表2给出了相应于 A、B、C、D 四类插值点的插值函数的准确值 FA 和各类误差的平均绝对值。由表2看出:

表2 插值函数的准确值FA及误差EL、EH1—EH3和EHA的平均绝对值/m

	FA	EL	EH1	EH2	EH3	EHA	平均*
A	255.7	0.7197	0.3596	0.1634	0.0673	0.0588	0.2738
B	258.9	0.9999	0.4980	0.2416	0.1157	0.1043	0.3919
C	255.7	0.7316	0.3599	0.1635	0.0672	0.0587	0.2762
D	262.4	0.7018	0.2212	0.1158	0.0678	0.0598	0.2333
平均	258.2	0.7883	0.3597	0.1711	0.0795	0.0704	

* 为EL、EH1、EH2、EH3、EHA的平均

(1)各种插值方案都有相当不错的精度。各方案误差的平均绝对值不超过1m。相对于约258m的平均准确值,其相对误差都在0.4%以下。

(2)在计算的3种插值方案中,L插值的误差最大。对A、B、C、D四类插值点平均,EL约为0.8m,为其他各类误差的2—10倍。根据下面对各方案插值余项的估计,该结论是正确的。

(3)H方案的各种误差中,以EH1最大,它约为EH2的2倍,为EH3的4.5倍,EH3最小,它与EHA已接近了。

(4)对A、B、C、D四类插值点而言,A类点和C类点的插值误差非常接近(见表2最右列),它们对各类误差的平均分别为0.274和0.276m;B类点的插值误差最大,为0.392m。即位于两格点之间中部的插值点(见图1)的插值误差较大,而靠近格点的插值点误差较小,后者约为前者的70%。以下对插值余项的估计表明这一结果是正确的。D类插值点误差最小,这与它只计算*i*=1, ..., 29,而不包含*i*=0的端点值有关。

4.2 误差大小的理论估计

在小区间 $[x_i, x_{i+1}]$ 内,H方案的插值余项为:

$$R_H(x) = \frac{f^{(4)}(\xi)}{4!} (x - x_i)^2 (x - x_{i+1})^2, \quad x_i < \xi < x_{i+1} \quad (12)$$

对A、C两类插值点均有:

$$R_{HA} = \frac{f^{(4)}(\xi)}{4!} \times 0.25^2 \times 0.75^2$$

$$\approx 0.0015f^{(4)}(\xi) \quad (13)$$

对B类插值点则为:

$$R_{HB} = \frac{f^{(4)}(\xi)}{4!} \times 0.5^2 \times 0.5^2 \approx 0.0026f^{(4)}(\xi) \quad (14)$$

Lagrangian 3次多项式的插值余项为:

$$R_L(x) = \frac{f^{(4)}(\xi)}{4!} (x - x_{i-1})(x - x_i)(x - x_{i+1})(x - x_{i+2}) \quad (15)$$

对A、C类插值点有

$$R_{LA}(x) = \frac{f^{(4)}(\xi)}{4!} \times 1.25 \times 0.25 \times 0.75 \times 1.75 \approx 0.0171f^{(4)}(\xi) \quad (16)$$

对B类插值点,则

$$R_{LB}(x) = \frac{f^{(4)}(\xi)}{4!} \times 1.5 \times 0.5 \times 0.5 \times 1.5 \approx 0.0234f^{(4)}(\xi) \quad (17)$$

由以上对插值余项的定量估计可见:

①A、C类插值点的插值误差相同且较小,这对H插值和L插值都成立。

②B类插值点误差较大,它比A、C类插值点的误差约大40%—70%。

③比较式(13)、(16);式(14)、(17)可见,对A、B、C各类插值点,L插值的误差约为H插值的10倍。当然,对于EH1、EH2的情况,由于边界值的影响,实际与EL之差没有这样大(见表2最下行)。

4.3 误差的离散程度

数组的均方根RMS与平均绝对值AAV之差值的大小能反映数组内样本的离散程度。例如,对于(1,9)和(5,5)两组数,它们的AAV值都是5.0,但RMS值分别为6.4和5.0。因此我们计算了各插值方案误差的RMS—AAV(用RA表示)值,见表3。表3表明,L插值虽然误差最大(见表2),但它的RA值小,即误差变动的振幅小。而在几种H插值计算中,误差的RMS—AAV之值大致与它们各自的平均绝对值AAV(见表2最下行)的大小成正比,故表3中平均而言是EH1的相应值最大,EH2次之。

表3 各方案误差的 RA/m 值

	EL	EH1	EH2	EH3	EHA	平均
A	0.1321	0.5973	0.1851	0.0073	0.0054	0.1854
B	0.0871	0.7300	0.2124	0.0108	0.0098	0.2100
C	0.0694	0.5973	0.1853	0.0074	0.0056	0.1730
D	0.0573	0.3132	0.0922	0.0075	0.0053	0.0951
平均	0.0865	0.5595	0.1688	0.0083	0.0065	

4.4 边界上的误差

计算的空间微商 m_i 在边界上的误差导致了 H 插值函数在边界上也有大的歪曲,这是 H 插值和 S 插值误差增长的主要原因。表 4 给出了 H 插值函数在边界上的准确值及误差的大小。表 4 与表 2 比较我们看到, EH1 和 EH2 在边界上都高出平均误差很多。对于 A、B、C 三类插值点, EH1 在 $i=0$ 和 29 两端点的误差约为整个区域平均值(见表 2 下行)的 7—12 倍; EH2 为 5—10 倍; EH3 由于采用了与准确解一致的周期性边界条件,故在边界上根本没有出现异常的误差。

表4 H 插值在边界上的准确值及误差/m

	i	准确值	EH1	EH2	EH3	EHA
A	0	62.57	4.38	1.55	-0.05	-0.03
	29	181.60	-2.43	-0.86	-0.02	-0.03
B	0	123.61	4.53	1.60	-0.05	-0.05
	29	-123.61	-4.53	-1.60	-0.06	-0.05
C	0	181.60	2.43	0.87	-0.02	-0.03
	29	-62.58	-4.38	-1.55	0.05	0.02
D	1	181.60	2.75	0.92	-0.02	-0.03
	29	-282.85	0.96	0.47	-0.09	-0.07

5 小结

通过以上数值计算和分析得到:

5.1 由于我们引进了三种精度较高的边界条件,由 Spline 函数方法计算的一阶空间微商 $m = \frac{dF}{dt}$ 都能达到一定的精度,但其中以采用周期性边界条件的误差最小。该误差的各项统计特征量比用二阶、三阶单向差外插边界值的相应误差要小 1—2 个量级;二阶单向差边条件导致的 m 误差比三阶单向差的相应结果要大一倍以上。

5.2 L、S、H 三种插值函数的最大相对误差均小于 4%,说明它们都具有相当好的精度。

其中,当上述改进的边界条件引入 H 和 S 方案的计算以后, L 方案的误差最大,但误差的离散程度小。对 A、B、C、D 四类插值点平均而言, EL 比 EH1 大一倍以上,比 EH2 大近 4 倍,比 EH3 大一个量级。

5.3 对于不同相对位置的 A、B、C、D 四类插值点,位于两格点中部的 B 类插值点误差较大;而位于靠近格点的 A、C 类插值点的误差较小,约为前者的 70%。以上结果均与对插值余项的理论分析一致。

5.4 S 插值与 H 插值实质相同。在 H 插值中,以采用周期性边条件所得插值函数的误差最小;采用二阶单向差边条件的误差值最大且分布也最为离散。它比三阶单向差边条件的相应误差约大一倍,比周期性边条件的相应误差约大 3.5 倍。

5.5 H 插值中,单向差边条件导致的误差主要表现在边界上。此时插值函数在边界上的误差可达整个区域平均值的 7—10 倍。周期性边条件对应的插值函数在边界上无明显误差。

参考文献

- 1 Mahrer, Y. and R. A. Pielke. A test of an upstream Spline interpolation technique for the advective terms in a numerical mesoscale model. Mon. Wea. Rev., 1978, 106: 818—830.
- 2 廖洞贤, 王两铭. 数值天气预报原理及其应用. 北京: 气象出版社, 1986: 237—249.
- 3 郑良杰. 二维三次样条函数在正压移动套网格数值预报模式中的应用试验. 数值天气预报文集. 北京: 气象出版社, 1984: 97—109.
- 4 Sandström, A. and T. Elvius. Numerical Method used in Atmospheric Models. II, Chapter 7, GARP (17), 1979.
- 5 徐萃薇编. 计算方法引论. 北京: 高等教育出版社, 1985: 13—53.
- 6 游性恬, 张兴旺. 数值天气预报基础. 北京: 气象出版社, 1992: 171—174.
- 7 Ahlberg, J. H., et al. The Theory of Splines and Their Applications Mathematics in Science and Engineering. Vol. 38. New York and London: Academic Press, 1967.

(下转第 40 页)

(上接第 7 页)

Experimental Comparisons Among Several Numerical Interpolation Methods

You Xingtian

(Beijing Meteorological College, 100081)

Abstract

Three higher-accurate boundary conditions is introduced. The first-order spacial differential is calculated by using the Spline function, and a set of simpler calculating formulation of the Hermite method is presented. Comparisons of the interpolation errors among the Lagrangian, the Spline and the Hermite schemes for different locations of the interpolation points are made. The results show that although these schemes have given accuracies to some extent, the error of the Hermite method with periodic boundary condition is around one fourth or a half of the errors of other boundary conditions, and is about one order of magnitude smaller than that of the Lagrangian technique.

Key Words: interpolation boundary condition error