

海洋资料整理中一种单调分段插值方法的应用

黄建东 郑逢中 彭荔红

(厦门大学环境科学研究中心, 厦门, 361005)

摘要 以三次 Hermite 插值函数和另二类初等形式的函数为基础所构造的一种混合型分段插值函数进行海洋资料整理中的插值和绘线处理, 结果更加合理。特别对具有累加或统计性质的数据, 不但给出光滑的插值函数及其图形, 而且还能给出大部分光滑的导函数图形。本文同时介绍了上述方法在不同专业领域中的一些应用实例。

关键词 资料处理 插值 单调光滑 海洋科学

对具有单调性质的离散数据进行插值与绘线处理, 需构造满足单调性且连续光滑的插值函数。已有文献对单调的三次样条、二次有理样条等插值函数作了讨论^[1-4]。这些单调插值函数, 有的要求型值点之间满足一定关系而无法直接应用于实际数据^[1,2]; 有的在处理变化特别剧烈或缓慢的一类数据时有良效, 但未能保证对任意给定的数据都达到一阶光滑^[3]; 有的由于区间上所有节点处的一阶导数都不是由特定方程组求出从而未能达到应有的光滑性^[4]。

对于一类累加型单调数据, 插值函数一阶以上的光滑性具有十分重要的意义。本文介绍一种以两点三次 Hermite 插值函数和其它两类初等形式的函数为基础构造的分段插值函数的应用。这种插值函数在保证单调性和一阶光滑性的基础上, 对于许多统计类数据, 在其所在区间的大部分中, 还使子区间内及节点上具有二阶光滑性。

1 分段插值函数概述

给定数据点列 $\{(x_i, y_i)\}_{i=0}^n, a = x_0 < x_1 < \dots < x_n = b, y_i < y_{i+1} (i=0, 1, \dots, n-1)$ 。需在 $[a, b]$ 上构造一连续的插值函数 $F(x)$, 满足单调性和较好的光滑性。记 $\Delta x_i = x_{i+1} - x_i, \Delta y_i = y_{i+1} - y_i, h_i = \Delta y_i / \Delta x_i; m_i$ 为 (x_i, y_i) 处的一阶导数值 $(i=0, 1, \dots, n)$; $F(x)$ 为 $[a, b]$ 上欲求取的插值函数, $F(x) = F_i(x), x \in [x_i, x_{i+1}]$ 。

首先, 在 $[a, b]$ 上求出三次样条插值函数 $S(x)$ ^[5], $S(x) = S_i(x), x \in [x_i, x_{i+1}]$ 。边界条件依 $[a, b]$ 两端点及其附近的型值取定。特别对于统计类数据, 在原有或外推得到的边界点处, 可取 $m_0 = 0$ 和 $m_n = 0$ ^[6]。由于实际数据并非都能满足三次样条插值函数的单调性条件, 故而存在这样的点 $x, x \in [x_i, x_{i+1}]$, 使 $S'(x) < 0$ 。记这样的子区间为 Δ 。在 $[x, x_{i+1}] \subset \Delta$ 上, 当 $[x_{i-1}, x_i] \subset \Delta$, 则取

本文于 1995 年 1 月 29 日收到。

$$m_i = h_{i-1} [\Delta x_i / (\Delta x_i + \Delta x_{i-1}) + \Delta y_i / (\Delta y_i + \Delta y_{i-1})] / 2 + h_i [\Delta x_{i-1} / (\Delta x_i + \Delta x_{i-1}) + \Delta y_{i-1} / (\Delta y_i + \Delta y_{i-1})] / 2 \quad (1)$$

否则,取 $m_i = S'_{i-1}(x_i)$ 。这样取定的 m_i 都满足 $m_i > 0$, 根据 m_i, h_i, m_{i+1} 三者的关系, 可考虑如下两种函数形式作为 $[x_i, x_{i+1}]$ 上的插值函数。

在 $[x_i, x_{i+1}] \subset \Delta$ 上, 当 m_i, h_i, m_{i+1} 顺次增大或减少时, 用形如

$$\lambda(x-\xi)^2 + \mu(y-\eta)^2 = \lambda \cdot \mu \quad (2)$$

的函数作为插值函数。以该子区间两端的型值和按以上所述给定的 m_i, m_{i+1} 值, 可得到由 4 个方程构成的方程组。此方程组当 $d = m_i + m_{i+1} - 2h_i \neq 0$ 时有唯一解。当 $d = 0$ 时, 对 $[x_i, x_{i+1}]$ 上的两点三次 Hermite 插值多项式^[5], 有 $H'(x) > 0$ 。故此时可直接用 $H(x)$ 作为 $[x_i, x_{i+1}]$ 上的插值函数且在该子区间两端其一阶导数连续。

对于 m_i, h_i 和 m_{i+1} 三者间不呈量值依次变化的情形, 考虑

$$y(x) = \lambda(x-\xi)^\mu + \eta \quad (3)$$

形式的函数。特别是在 $[x_i, x_{i+1}]$ 上, 分别对 $[x_i, \xi]$ 和 $[\xi, x_{i+1}]$ 求出如式(3)的插值函数。由 $[x_i, x_{i+1}]$ 两端型值和给定的 m_i, m_{i+1} 以及由两点三次 Hermite 插值多项式得到的两端二阶导数, 由式(3)可得到 6 个方程构成的方程组。经数据和变量的变换, 该方程组若有符合要求的解, 则得一组待定系数值; 若解不符合要求或无解, 则可适当指定 (ξ_i, η_i) , $\xi_i \in [x_i, x_{i+1}]$, $\eta_i \in [y_i, y_{i+1}]$ 。此时根据型值和一阶导数值经式(3)构成的方程组, 可求得所需待定系数值, 且解存在、唯一。

在 $[a, b]$ 的各子区间上, 若 $[x_i, x_{i+1}] \subset \Delta$, 则令 $F_i(x) = y_i(x)$ 或 $F_i(x) = H_i(x)$, 否则令 $F_i(x) = S_i(x)$, 便得到 $[a, b]$ 上的分段插值函数 $F(x)$ 。由 $F(x)$ 的构造知, ① $F'(x) > 0, x \in [a, b]$; ② 在 $F_i(x) = S_i(x)$ 的连续子区间段内部包括节点上, 都有 $F''(x)$ 连续; ③ 在其余子区间的节点上, $F'(x)$ 连续^①。

2 应用与实例

在数据区间 $[a, b]$ 上, 若单以三次样条函数作为插值函数, 则可能出现插值函数的摆动振荡(图 1a₁, b₁), 从而使曲线不满足单调变化类数据本身的性质; 而按前述一些方法构造的单调插值函数, 当 $F'(x)$ 不连续时^[3], 频率分布曲线将出现间断; 而每个节点上 $F''(x)$ 的不连续^[4], 则使频率分布曲线出现许多不光滑的转折点(图 1a₂, b₂)。应用上希望有较光滑的频率分布曲线的要求超出了上述方法的适用范围。本文 $F(x)$ 的构造对曲线摆动振荡和光滑性问题都作了较合理的处理, 在实际数据处理的应用中获得了更好的效果。

单调型数据中, 有类似于水温沿水深变化的一类数据; 还有一类是具有统计性质的数据, 如波浪的波高累积率, 海滩和海底沉积物的粒度分布, 海洋生物类的统计数据等等, 其中大量是服从或近似于正态分布的。依文献[6]所述并经大量插值绘线实例验证, 对于服从正态分布的统计类数据, 可在其数据区间 $[a, b]$ 的大部分中用三次样条插值函数作为插值函数并满足单调性。实际数据处理还表明, 对于来自多个正态总体的数据, 以及不服从正态分布的数据, 三次样条函数仍可在数据区间的相当大的部分中满足单调性。这样, 在 $[a, b]$ 的大部分里插值函数 $F(x)$ 的导函数便具有一阶导数的连续性, 也就使频率分布曲线的大部分具有一阶光滑的性

① 黄建东. 单调光滑插值函数的一种构造及其在统计类数据处理中的应用. 1995.

质。

下面给出海洋研究中不同专业领域的几个实例。所有数据均由实测获得。应用插值函数 $F(x)$,除了可以插值求数据区间上任一点的函数值以及求算任意百分位值以外,例子中均给出了 $F(x)$ 导函数的图形,即在 $[a, b]$ 的大部分中连续光滑的频率分布曲线。

2.1 沉积物粒度分布

表 1 是两个海滩沉积物的粒度分布数据。据之所绘累积频率曲线和频率分布曲线如图 $1a_3, b_3$ 。

表 1 海滩沉积物粒度分布数据

Tab. 1 Grain size distribution data (cumulative frequency) of two sediment samples from a beach

粒级(Φ 值)		-1.68	-1.00	0.00	0.74	1.49	2.74	3.37	4.16
累计频率 (%)	a 组	0.0	0.32	1.78	11.55	39.90	98.30	99.82	100.00
	b 组	0.76	10.37	20.44	33.68	49.08	96.54	99.53	100.00

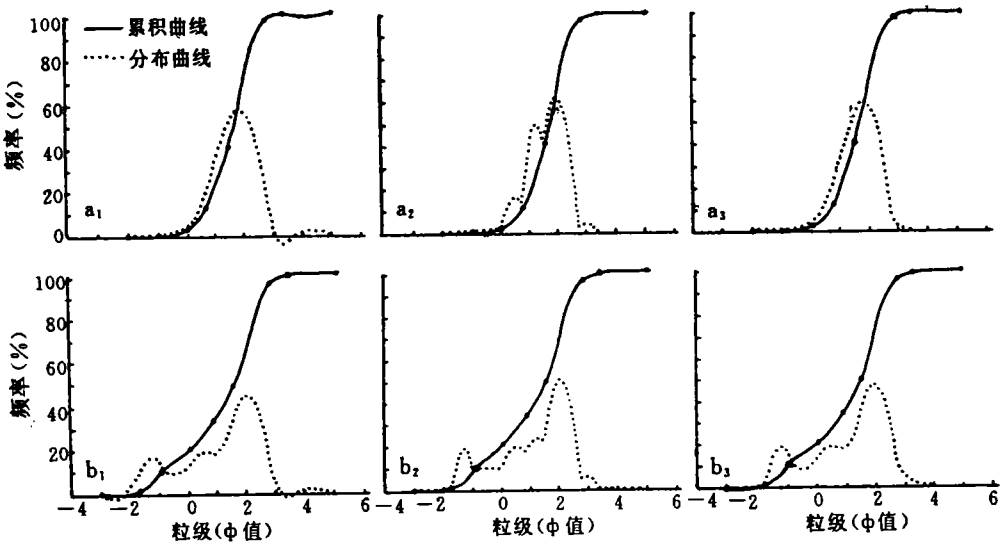


图 1 海滩沉积物粒度累积频率曲线和频率分布曲线

Fig. 1 Cumulative and frequency curves of sediment samples from a beach

a_1, b_1 : 三次样条函数; a_2, b_2 : 文献[4]方法; a_3, b_3 : 本文方法

由图、表可见, a 组数据反映较单一的物源及动力环境,而 b 组则反映较多样的物源或动力环境,或二者的结合。由频率分布曲线可直观地了解沉积物特性;除了总的状况外,尚可据曲线的单峰、多峰,各峰的低、峰的位置等等获得关于沉积物搬运动力、沿程变化和沉积环境等的有价值的信息;较之概率累积曲线,在详细了解沉积物特性及其环境特征方面有独到的作用。

图 1 还给出了同一组数据利用其他方法所得结果(图 $1a_1, b_1, a_2, b_2$),由三者对比可见本文方法(图 $1a_3, b_3$)所具有的更合理的处理效果。

2.2 海洋环境生态实验数据

表 2 为红树植物秋茄幼苗在相同试验条件下,受不同浓度重金属镉影响,生长 30d 后幼苗

高度的两组数据。a 组的幼苗中有一部分为未受镉影响(对照组);另一部分分二组,分别受两种较低的不同浓度镉影响。而 b 组则在较高的三种镉浓度下生长。图 2 为以上数据的累积频率曲线和频率分布曲线。图 2a 中反映了三部分次总体,即反映出未受镉污染和受两种较低浓度镉污染的三部分群体的生长高度情况,每个峰的位置大致反映了各该部分群体的平均生长高度。图 2b 则表现出某种一致性,幼苗高度平均值向低的方向移动。即污染程度较重的群体其苗高明显低于污染较轻的群体。且浓度超过一定阈限后,浓度在一定范围内的增高对幼苗高度的影响似表现不明显。

表 2 红树植物秋茄幼苗在不同浓度镉影响下生长 30d 后幼苗高度统计
Tab. 2 Cumulative frequency of seedling height distribution of two groups of Mangrove *Kandelia candel* suffered from Cd pollution for 30 d

幼苗高度(cm)		0	3	6	9	12	15	18
累积频率 (%)	a 组	0.00	2.22	35.56	66.67	93.33	97.78	100.00
	b 组	0.00	15.56	51.11	82.22	100.00	—	—

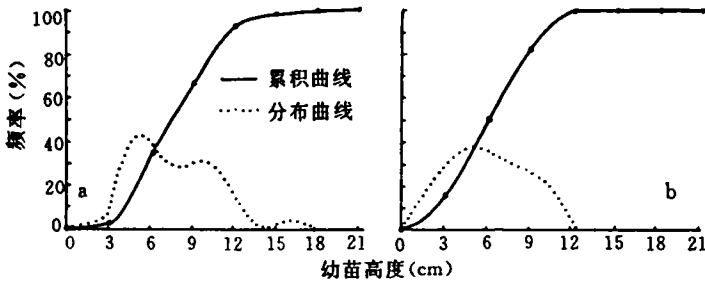


图 2 秋茄幼苗在不同浓度镉影响下生长 30d 后幼苗高度累积频率曲线和频率分布曲线

Fig. 2 Cumulative and frequency curves of two groups of *Kandelia candel* suffered from Cd pollution for 30d

2.3 波浪统计数据

表 3 是波浪相对波高累积率的两组数据(据文献[7],经变换)。图 3 为对这些数据插值绘线的结果。其中,可以看到数据所反映的有可能是参数略有不同的若干列波浪叠加的情况,尤以 b 组较明显。

表 3 波浪相对波高累积率
Tab. 3 Two groups of cumulative frequency data of relative wave height

相对波高 (m)	a 组	0.00	0.51	0.75	0.97	1.20	1.33	1.54	1.72	2.01	2.36
	b 组	0.00	0.49	0.72	0.96	1.21	1.34	1.59	1.77	2.10	2.48
累积频率 (%)		0.0	10.0	30.0	50.0	70.0	80.0	90.0	95.0	99.0	99.9

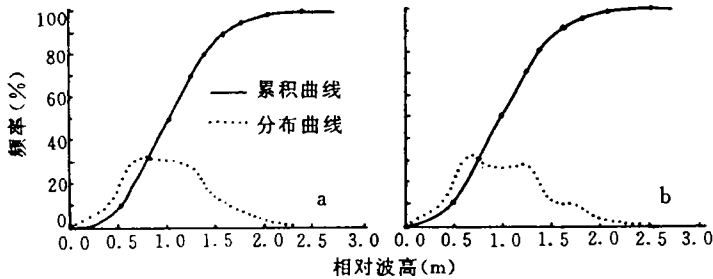


图 3 波浪相对波高累积频率曲线和频率分布曲线

Fig. 3 Cumulative and frequency curves of relative wave height

3 结语

从上述海洋沉积、水文和生态等方面的数据处理实例可以看到,利用插值函数 $F(x)$ 的单调光滑性得到的插值和绘线的计算机处理结果,特别是由 $F''(x)$ 连续的性质得到的频率分布光滑曲线,对于在更符合客观实际的基础上了解研究对象若干本质性的特征,从而获得规律性的认识和判断,具有重要的意义。

应用插值函数 $F(x)$,已对数百例各种实际数据进行了插值、绘线处理,获得了较其它一些方法更为合理的结果。其中大部分得到了满意的 $F'(x)$ 的图形。对于非统计类的单调数据,也可作出质量好的处理。这种情况下不要求绘出 $F'(x)$ 的图形,故 $F(x)$ 达一阶光滑一般即可满足需要。

混合型分段单调插值函数 $F(x)$ 既满足了对实际单调数据处理的需要而无需特定的数据点组合条件,又具有尽可能好的光滑性,这种性质使之在海洋观测资料单调型数据的插值及绘线处理中,表现出良好的适用性和实用性。

参考文献

- 1 文涛. 单调光滑函数的样条插值. 计算数学, 1981, (2): 143~151
- 2 黄达人. 单调三次样条的一个充分条件. 计算数学, 1982, (2): 214~217
- 3 黄龙生. 单调堡型样条及其在海洋资料整理中的应用. 海洋通报, 1992, 11(1): 77~83
- 4 张宝琳. 逐段单调三次样条插值. 数值计算与计算机应用, 1983, (3): 157~162
- 5 易大义等编. 数值方法. 杭州: 浙江科学技术出版社, 1984. 34~52
- 6 黄建东. 两种概率分布函数的样条逼近. 厦门大学学报(自然科学版), 1988, 27(2): 234~236
- 7 文圣常, 余宙文编著. 海浪理论与计算原理. 北京: 科学出版社, 1984. 176~177

Application of a sort of monotonic piecewise interpolation method in marine data Processing

Huang Jiandong, Zheng Fengzhong and Peng Lihong

(*Research Centre of Environmental Sciences, Xiamen University, Xiamen, 361005*)

Abstract

A sort of multicomponent piecewise interpolation function, which is composed of the cubic Hermite Interpolation Function and two other sorts of the elementary functions, possesses the properties as the following: 1. monotonicity in the whole interval $[a, b]$; 2. continuous second derivative in those subintervals (esp. at the knots) in which the Cubic Spline Function can be assigned to be the interpolation function; 3. continuous first derivative at the rest knots. The application of this sort of interpolation function is introduced in this paper, with its satisfied performance. Especially, for the data with the characteristics of statistics and cumulation, not only a monotone and smooth interpolation function and its curve can be derived by this method, but also a smooth derivative function and its curve in most subintervals of the interval $[a, b]$, by means of computer. Several examples of the application of this method in data processing in different disciplines of the marine sciences are illustrated.

KEYWORDS Data processing, interpolation, monotonic and continuous, marine sciences