www.scichina.com

tech.scichina.com



基于参数函数曲面的参数率定方法

包为民®®,张小琴®®*,赵丽平®®

- ① 河海大学水文水资源与水利工程科学国家重点实验室, 南京 210098;
- ② 河海大学水文水资源学院, 南京 210098
- * E-mail: zxqin403@163.com

收稿日期: 2012-10-09; 接受日期: 2013-04-10

国家自然科学基金(批准号: 51279057)、国家自然科学基金重大项目(批准号: 51190090 和 51190091)、高等学校博士学科点专项科研基金(批准号: 20120094120018)和中央高校基本科研业务费专项资金(批准号: 2012B00214)资助

摘要 针对参数估计过程中普遍存在的问题,即不同初始值获得不唯一的参数估计结果,深入分析了误差平方和目标函数的结构及其对参数估计提供的信息特性,发现了目前参数估计方法所存在的本质性问题: 1) 误差平方和目标函数为参数估计提供的信息通常是不合理的甚至是错误的; 2) 误差平方和目标函数曲面比参数函数的曲面更复杂,参数优值不适合在目标函数曲面上寻找,而应在参数函数曲面上寻找。由此,本文提出了样本截痕的概念,分析论证了截痕交点唯一性定理(即参数优值的唯一性定理)。根据参数函数、曲面信息特点与泰勒级数展开方法,提出了基于参数函数曲面的截痕相交参数估计方法。经理论论证和参数估计应用检验,样本截痕相交参数估计方法避免了现有参数自动率定的问题,方法直观简单,而且证明了方法的收敛性和应用效果。

关键词

参数估计 目标函数型 唯一性定 植痕相交 泰勒级数

1 引言

参数估计涉及自然科学和社会科学的各个领域, 科学研究者对其关注程度、投入的科研力量和获得的 成果数量都是相当巨大的.模型的成功应用依赖于 参数估计,由于参数人工率定费时费力,参数自动估 计方法得到了大量研究.优化算法和目标函数是参 数自动估计的关键.参数估计除物理模型和一些能 用经验方法估计的参数外,都需要借助于目标函数 获取参数估计的信息,误差绝对值的正整数次方是 常用的目标函数公式,其中误差平方和应用最广.

传统的估计方法,如牛顿法^[1]、下降梯度法、Rosenbrock 法^[2]和 Powel 法^[3],不能用于解决存在多个局部优值的问题.由于水文模型大多数是非线性

的,模型响应面是多峰的,参数空间里存在若干局部优值点,上述传统方法难以越过不连续目标函数曲面上的局部优值.此外,参数关联性、不唯一性(non-uniqueness)^[4,5]、不适定性(ill-posedness)^[6,7]、参数等效性(equifinality)^[8]等诸多问题也会影响传统方法,导致由不同参数初值得到的参数估计结果不唯一,但又无法证明模型是存在局部优值还是方法不当所致^[9,10].随着计算能力的提高,一些以寻找全局唯一优值为出发点的复杂优化方法得到快速发展,例如,基因遗传算法^[11,12]、蚁群算法^[13,14]、粒子群方法^[15,16]、人工神经网络^[17]和 SCE-UA 方法^[18,19]等.这些以确定性寻找信息与随机相结合的现代优化方法,试图避免传统方法存在的问题,以得到全局最优的、唯一的参数估计.但该类方法同样无法证明是否存在局

英文版发表信息: Bao W M, Zhang X Q, Zhao L P. Parameter estimation method based on parameter function surface. Sci China Tech Sci, 2012, 56: 1485-1498, doi: 10.1007/s11431-013-5224-3

部优值,也无法证明所获得的参数值是否为全局最优,甚至连参数估计方法本身的收敛性也常难以证明,只在概率意义上分析能收敛于全局优值^[20-22].但随机意义上的分析显然不可能回答、更不可能解决前面所述的确定性范畴的问题.

目标函数曲面通常是不光滑的、多模式的、不连 续的, 采用不同的目标函数会得到不同的参数估计 值[23, 24]. 要提出一种新的具有较强能力的参数率定 方法必须要识别这种内在的多目标本质[25]. 本文作 者[26]在20世纪80年代的博士论文研究中提出,现有 参数估计方法的根本问题在于目标函数的黑箱子结 构和基于目标函数的信息利用方法, 如何构建新的 目标函数或如何把黑箱子的目标函数物理化, 使得 不同参数估计能提取相应的信息, 是解决参数估计 的关键. 研究中先后提出了多结构目标函数构建与信 息利用[27]、物理概念模型参数的概念化率定方法[28]、 抗差性目标函数构建与信息利用[29~34]、定义于参数灵 敏度域的目标函数与信息利用[35~37]、非线性模型参数 的线性化估计方法[38]、全局优值跨峰率定方法[39]等. 这些研究只是对传统参数估计方法有所改进, 也不 能彻底解决现有参数估计方法中存在的问题.

本文从误差平方和目标函数的结构、对参数估计能提供的信息分析入手,发现了目前参数估计方法的根本问题所在,提出了截痕交点的概念、截痕交点唯一性定理(即参数优值的唯一性定理).针对参数估计在目标函数曲面上寻找方法的问题,提出一种直接在参数函数曲面上寻找最优值的方法.本文通过分析泰勒级数的特性,提出了截痕相交参数估计方法,证明了方法的收敛性,检验了方法使用的效果.

2 误差平方和目标函数问题

目标函数是参数估计最重要的信息来源,这信息决定了所估计参数的效果.目前参数估计绝大多数方法以误差平方和目标函数为基础,论文研究发现,误差平方和目标函数信息对参数估计存在诸多问题,有些情况下甚至会提供错误的参数估计信息.

2.1 单样本信息的不合理性

误差平方和目标函数以单样本信息为基础,其 为参数估计提供的信息是方法成败的关键.为简单 又不失一般性,这里以单参数函数模型为例进行分 析. 一般形式的单参数模型为

$$y = f(\theta, x), \tag{1}$$

其中 y 为模型输出(应变量), x 为模型输入(自变量), θ 为模型参数. 对于模型的一个样本(x_1,y_1), 构造误差平方和目标函数 F_1 如下:

$$F_1 = \left(y_1 - f\left(\theta, x_1\right)\right)^2. \tag{2}$$

(2)式是单样本为参数估计提供的误差平方信息, 这里简称为样本信息.

为清楚地说明单参数模型样本信息的不合理性, 这里举一个具体函数来说明.

$$y = 2 + (x\theta)^2 + \sin(2\pi x\theta). \tag{3}$$

对于(3)式函数, 当参数真值为 θ =0.3 时, 取两个样本(1, f(0.3, 1))和(2, f(0.3, 2)), 其相应的样本误差平方目标函数为

$$F_1(\theta) = \left(f\left(0.3, 1\right) - \left(2 + \theta^2 + \sin\left(2\pi\theta\right)\right) \right)^2, \tag{4}$$

$$F_2(\theta) = \left(f\left(0.3, 2\right) - \left(2 + \left(2\theta\right)^2 + \sin\left(4\pi\theta\right)\right) \right)^2. \tag{5}$$

(4)和(5)式都是参数 θ 的函数,目标函数值随参数值的变化过程见图 1.图 1中横坐标是参数,纵坐标是样本误差平方目标函数值.

从图 1 看出,样本 1 相应目标函数存在 10 个局部最小值,而样本 2 为 9 个,这些局部范围内的最小值(又称局部优值)具体数值见表 1.表 1 中 θ_{1min} 和 F_{1min} 分别为样本 1 误差平方目标函数的参数局部优值和目标函数值, θ_{2min} 和 F_{2min} 分别为样本 2 误差平方目标函数的参数局部优值和目标函数的参数局部优值和目标函数值.

从图 1 和表 1 看出,不同样本信息显示的参数局部优值位置和数量都不同,每增加一个样本理论上就有可能增加一组局部优值,这显然是不合理的.

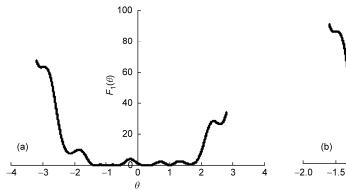
2.2 不同样本组合信息的不合理性

目标函数由不同样本误差平方累加构成,对于两组不同的样本,希望提供参数优值估计信息能引向相同的局部优值,但事实却不是这样.如对于函数(3)式的两组样本为

$$(x_i, f(0.3, x_i)), x_i = 1, 2, 10,$$
 (6)

$$(x_i, f(0.3, x_i)), x_i = 1, 2, 3, \dots, 10.$$
 (7)

其相应的样本误差平方和目标函数为



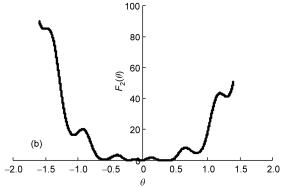


图 1 目标函数值随参数的变化过程

(a) x=1 样本误差平方目标函数; (b) x=2 样本误差平方目标函数

表 1 不同样本的局部优值

Number of local optima	$ heta_{ m 1min}$	$F_{1\min}$	$ heta_{2 ext{min}}$	$F_{2\min}$
1	-3.309	63.232	-1.5205	85.021
2	-2.13	7.6556	-1.066	16.287
3	-1.353	6.79×10^{-5}	-0.5935	0.5099
4	-0.99	3.36×10^{-6}	-0.2155	4.28×10^{-5}
5	-0.615	2.31×10^{-6}	-0.019	5.28×10^{-5}
6	0.228	1.93×10^{-6}	0.3	1.51×10^{-31}
7	0.3	1.73×10^{-33}	0.41	2.15×10^{-5}
8	1.005	1.44×10^{-7}	0.83	4.5756
9	1.662	0.7572	1.298	40.956
10	2.595	26.3259		

$$F_1(\theta) = \sum_{i=1}^{3} (f(0.3, x_i) - (2 + (x_i \theta)^2 + \sin(2\pi x_i \theta))^2, (8)$$

$$F_2(\theta) = \sum_{i=1}^{10} (f(0.3, x_i) - (2 + (x_i \theta)^2 + \sin(2\pi x_i \theta))^2).$$
 (9)

(8)和(9)式是参数 θ 的函数,二者的局部最小值如表 2 所示,(8)式相应目标函数存在 7 个局部最小值,而(9)式为 5 个,具体数值见表 2.

从表 2 可知,不同样本组合时,目标函数信息显示的参数局部优值位置和数量都不同,这也是不合理的.

2.3 多参数模型样本信息的不合理性

对于下面所示最简单的两参数模型:

$$y = (x\theta + 1)^2 + (\beta + x^2)^2$$
. (10)

如果真参数 θ =3, β =1, 则对于样本(x=1, y=20), 参数满足如下关系(11)式都是其优值, 即存在无穷多个优值.

$$20 = (\theta + 1)^{2} + (\beta + 1)^{2}. \tag{11}$$

类似地可以分析,参数两个及其以上的多参数模型,其参数模型样本信息显示的参数优值都为无穷多个. 这样的信息为误差平方和目标函数增加了复杂性.

2.4 不同参数样本组合的错误性

当模型参数随着样本的不同有变化时,其样本组合就会提供错误的信息.如(3)式所示的单参数模型,如果参数随自变量不同而变化,如下式所示:

$$\begin{cases} \theta = 0.3, & x \le 10, \\ \theta = 0.1, & x > 10. \end{cases}$$
 (12)

则对于其参数值不同的样本(1, f(0.3, 10))和(11, f(0.1, 11))组合的目标函数为

表 2 不同样本组合的局部优值

Number of local optima	$ heta_{ m 1min}$	$F_{1 m min}$	$ heta_{2 ext{min}}$	$F_{2 m min}$
1	-0.319	3.6454	-0.28281	18.78
2	-0.185	28.4578	-0.09381	140.2
3	-0.078	58.53	0.03076	142.9
4	0.027	63.79	0.3	1.957×10^{-30}
5	0.132	40.665		
6	0.239	7.18		
7	0.3	6.27×10^{-30}		

$$F(\theta) = \left(f(0.3, 10) - \left(2 + (10\theta)^2 + \sin(20\pi\theta) \right) \right)^2 + \left(f(0.1, 11) - \left(2 + (11\theta)^2 + \sin(22\pi\theta) \right) \right)^2. (13)$$

当参数变化时, 其局部最小值如表 3.

从表 3 看出,局部优值中既不包含 0.1,也不包含 0.3.类似地可以分析,对任意变参数模型、不同参数样本的组合,所得结果一般不包含真值参数.显然根据此信息估计模型参数就不可能获得参数的真值,只能得出错误的参数估计结果.

2.5 误差平方和目标函数构建问题

几乎所有的参数率定都以误差平方和目标函数 为参数确定的信息依据. 误差平方和目标函数参数 率定方法, 通常涉及误差平方和目标函数构建和一 阶函数求导为零得其解这两步操作. 这些操作对于 线性参数没有任何问题, 在理论上和应用中都能严 格地证明其方法能得到唯一的最优值. 但线性参数 估计方法延伸到非线性模型参数的率定中, 发现存 在目标函数构建给非线性参数增加了不相关的局部 优值问题.

以误差平方和为信息基础的传统参数估计方法, 对(1)式函数构建误差平方和目标函数如下式:

$$\min_{\theta \in \mathfrak{R}} \left\{ F(\theta, x) = \sum_{i=1}^{L} \left(f(\theta, x_i) - y_i \right)^2 \right\},\tag{14}$$

其中 $\{x_i, y_i; i=1,2,\cdots,L\}$ 为确定参数的样本系列.

(14)式的平方和操作使线性参数的次数由 1 次增加为 2 次,又通过求导由 2 次降为 1 次,不改变参数次数和优值数量,所以这对于线性函数参数优化是没有任何问题的.但对于非线性参数函数就带来了问题,如果参数的方次大于等于 2,通过平方操作参数方次加倍,而求导一次只降低参数方次 1 次,所以就增加了参数解的个数.

为说明简单,这里以一般n次参数的非线性函数为例:

$$y = x + \theta x + (\theta x)^2 + \dots + (\theta x)^n, \quad n \ge 2.$$
 (15)

构建平方和目标函数:

表 3 变参数模型局部优值

$ heta_{ ext{min}}$	-0.2776	-0.1878	-0.0787	0.026	0.1329	0.2369
F_{\min}	51.88	27.12	54.96	63.5	40.7	26.19

$$F = \sum_{i=1}^{L} \left[x_i + \theta x_i + (\theta x_i)^2 + \dots + (\theta x_i)^n - y_i \right]^2.$$
 (16)

经过一阶导数操作(16)式变为

$$\frac{\mathrm{d}F}{\mathrm{d}\theta} = 2\sum_{i=1}^{L} \left[x_i + \theta x_i + \left(\theta x_i\right)^2 + \dots + \left(\theta x_i\right)^n - y_i \right] \times \left[x_i + 2x_i \theta x_i + \dots + nx_i \left(\theta x_i\right)^{n-1} \right]. \tag{17}$$

使参数解数量增加了(*n*-1)个,而且这增加的(*n*-1)个解不是函数本身的,是通过平方操作额外增加的.如果用这样的目标函数优选来确定参数,可能得到不相关的参数值,显然是没有意义的.

工程应用中有许多函数或模型参数是有物理意义的,用误差平方和最小目标函数方法会获得物理意义上不合理的结果,该类问题显然与这些额外解的存在有关.由于误差平方和目标函数存在如上分析的诸多问题,有必要思考新的参数估计方法和信息利用手段.

3 样本截痕信息对参数估计的作用与唯一 性定理

对于函数或模型的自变量取定值时,其函数值随参数而变化,形成定义在参数空间域上的函数曲面.此函数曲面的变化特性与可提取的参数估计信息,显然对参数估计是最重要和最直观的.为了便于分析和研究,先给出如下3个定义.

定义一:参数函数.对于一个给定的观测样本 (x_i, y_i) ,其自变量代入函数或模型得

$$y(\theta) = f(x_i, \theta), \tag{18}$$

其中 x_i 取观测的确定值, y 随参数 θ 变化, 把由(18)式 定义的函数称为参数函数.

定义二: 样本函数. 将观测样本 (x_i, y_i) 代入函数或模型得

$$y_i = f(x_i, \theta), \tag{19}$$

其中只有参数*θ*是变量,是定义于参数域的函数.不同的样本其相应定义于参数域的函数不同,这里把由特定观测样本相应的参数域函数式称为样本函数.

定义三: 样本截痕. 特定观测样本(x_i, y_i)与其参数函数产生交点(对单参数模型)或交线(对两参数模型)或截面(对3参数以上模型), 这里统称样本与参数函数的相交为样本截痕. 参数估计最希望的是直接

获得相交位置相应的参数值. 所以下面分析截痕交点的特性.

3.1 样本截痕信息显示较少局部优值

图 2 是函数(3)式的样本(12, f(0.3,12))与函数相交显示的交点信息. 图 2 中曲线是 x=12 时函数随参数的变化,与参数轴平行的直线为样本值 f(0.3,12)与曲线的横截线,图右边的交点反映了参数真值的位置. 从图 2 可看出,交点的数量远少于局部最小值数量.

3.2 样本截痕信息直接反映真参数位置

每个样本都与参数函数相交, 所以样本截痕信息都能直接反映真参数的位置. 表 4 是函数(3)式与不同样本截痕获得的交点信息.

从表 4 可知,虽然不同的样本与参数函数的交点不同,但都包含相同的真参数位置 0.3.

3.3 不同样本截痕相交交点的唯一性定理与参数 优值

从表 4 结果看,不同样本截痕交点不同,除真值 交点相同外,其他交点几乎均不相同,当样本增加到 一定数量时,所有样本具有相同的交点是唯一的,这

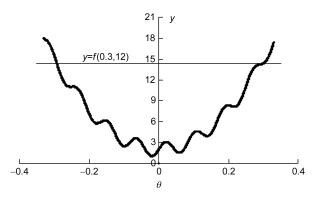


图 2 样本与函数的交点

表 4 函数与不同样本截痕交点

f(0.3,1)	f(0.3,2)	f(0.3,3)	f(0.3,5)	f(0.3,10)	f(0.3,11)	f(0.3,12)
-1.3543	-0.2159	-0.1655	-0.3	-0.3072	-0.3153	-0.2925
-0.9905	-0.0183	0.0119	0.2341	-0.3001	0.3	0.3
-0.6153	0.3	0.1685	0.3	-0.2903		
0.2269	0.4105	0.3	0.3585	0.3		
0.3						
1.0049						

就是真值参数所在的交点. 以两参数函数(10)式为例进一步说明这问题.

表 5 是不同样本在两参数函数曲面上获得的截痕曲线. 对于两参数函数,虽然其截痕交点同样有无穷多个,但这些截痕曲线间的交点是有限的,其在参数平面上的投影存在着数量不同的交点,但共同的交点是唯一的,即参数真值点(θ =3, β =1).

把 x=1, 2, 3 和 5 的截痕曲线分别展开:

$$\begin{cases} \theta^{2} + 2\theta + \beta^{2} + 2\beta = 18, \\ 4\theta^{2} + 4\theta + \beta^{2} + 8\beta = 57, \\ 9\theta^{2} + 6\theta + \beta^{2} + 18\beta = 118, \\ 25\theta^{2} + 10\theta + \beta^{2} + 50\beta = 306. \end{cases}$$
(20)

方程组系数阵的秩为 3, 简化为 3 个方程的方程 组为

$$\begin{cases} 2\theta + \beta^2 + (\theta^2 + 2\beta) = 18, \\ 4\theta + \beta^2 + 4(\theta^2 + 2\beta) = 57, \\ 6\theta + \beta^2 + 9(\theta^2 + 2\beta) = 118. \end{cases}$$
 (21)

(21)式中把 θ , β^2 和 θ^2 +2 β 作为 3 个由 θ 和 β 组成的新变量,可以直接解得

$$\theta = 3, \quad \beta^2 = 1, \quad \theta^2 + 2\beta = 11.$$
 (22)

进一步解得(θ=3, β=1)是(21)式的唯一解, 所以方程组(20)的交点唯一. 对于一般正整数指数函数的交点唯一性定理的证明见附录. 以上分析的"样本截痕曲线交点的唯一性"是非常有用的特性, 因为这就是参数估计所要寻找的参数优值位置. 要研究新的参数估计方法和信息利用手段, 自然就会想到利用该特性来构造参数估计所需的信息函数和相应参数估计方法.

4 样本截痕相交参数估计方法

4.1 参数函数与泰勒级数

从以上分析可知, 给定的自变量值不同, 相应的

表 5 样本截痕曲线

х	у	Curves of sample intersection	Shape
0	1	$1=\beta^2$	Point (3, 1)
1	20	$20=(\theta+1)^2+(\beta+1)^2$	round
2	74	$74 = (2\theta + 1)^2 + (\beta + 4)^2$	ellipse
3	200	$200=(3\theta+1)^2+(\beta+9)^2$	ellipse
5	932	$932=(5\theta+1)^2+(\beta+25)^2$	ellipse

参数域函数曲面是不同的,但其样本交点位置都包含了参数真值. 那么这函数曲面上有什么信息能直接将参数估计引向真值位置的方法是所希望寻求的. 对于给定自变量的样本函数曲面上任意点,如果函数在参数*θ*⁰⁾邻域内连续、存在任意阶的导数,有关于参数的泰勒级数展开:

$$f(x_{i},\theta) - f(x_{i},\theta^{(j)})$$

$$= \frac{\partial f(x_{i},\theta)}{\partial \theta} \bigg|_{\theta = \theta^{(j)}} (\theta - \theta^{(j)}) + \frac{1}{2} (\theta - \theta^{(j)})^{\mathsf{T}} \frac{\partial f^{2}(x_{i},\theta)}{\partial \theta^{2}} \bigg|_{\theta = \theta^{(j)}}$$

$$\times (\theta - \theta^{(j)}) + \cdots, \tag{23}$$

其中 x_i 为特定的自变量, $\theta^{(i)}$ 为第 j 步获得的参数估计值, $f(x_i,\theta^{(i)})$ 是函数用 $\theta^{(i)}$ 计算的值。 $f(x_i,\theta)$ 是真参数相应的函数值,如果函数没有任何误差,其值等于 y_i , θ 为要估计的参数.等号左边表达了函数的计算误差,

$$\left. \frac{\partial f(x_i, \theta)}{\partial \theta} \right|_{\theta = \theta^{(j)}}$$
 和 $\left. \frac{\partial f^2(x_i, \theta)}{\partial \theta^2} \right|_{\theta = \theta^{(j)}}$ 分别是函数在 $\theta^{(j)}$ 的一

阶和两阶导数,如果参数为两个以上的向量,则一阶导数为向量,两阶导数为矩阵.显然(23)式直接反映了误差(y_i - $f(x_i,\theta^{(i)})$)与参数(θ - $\theta^{(i)}$)ⁱ间的关系.

4.2 泰勒级数参数函数与参数估计

由于不同的样本,其样本函数曲面不同,所以(23)式只反映了不同样本函数曲面参数与误差间的关系.那么问题是如何用这不同样本函数曲面的关系来估计参数值.首先把(23)式表达为 *M* 阶泰勒级数的标量函数形式:

$$f(x_{i},\theta)$$

$$= \sum_{m_{1}}^{M} \cdots \sum_{m_{n}}^{M} \frac{\partial^{m_{1}}}{\partial \theta_{1}^{m_{1}}} \frac{\partial^{m_{2}}}{\partial \theta_{2}^{m_{2}}} \cdots \frac{\partial^{m_{n}}}{\partial \theta_{n}^{m_{n}}} \frac{f(x_{i},\theta)}{m_{1}! \cdots m_{n}!} \bigg|_{\theta=\theta^{(j)}}$$

$$\times \left(\theta_{1} - \theta_{1}^{(j)}\right)^{m_{1}} \cdots \left(\theta_{n} - \theta_{n}^{(j)}\right)^{m_{n}} + \xi_{i}. \tag{24}$$

由(24)式可知, *M* 次方的参数函数表达了曲面值的近似变化. 对于单参数函数(24)式变为

$$f(x_{i}, \theta) = f(x_{i}, \theta^{(j)}) + \frac{\mathrm{d}f(x_{i}, \theta)}{\mathrm{d}\theta}\Big|_{\theta = \theta^{(j)}} \times (\theta - \theta^{(j)}) + \dots + \frac{\mathrm{d}^{M}f(x_{i}, \theta)}{\mathrm{d}\theta^{M}}\Big|_{\theta = \theta^{(j)}} (\theta - \theta^{(j)})^{M} + \xi_{i}. \quad (25)$$

$$\xi^{(j)} = \begin{bmatrix} \xi_{1}^{(j)} \\ \xi_{2}^{(j)} \\ \vdots \end{bmatrix},$$

下面为简单以(25)式为例进行说明,将观测样本

系列代入(25)式得观测方程:

$$\begin{aligned} y_1 - f\left(x_1, \theta^{(j)}\right) \\ &= \frac{\mathrm{d}f\left(x_1, \theta\right)}{\mathrm{d}\theta} \bigg|_{\theta = \theta^{(j)}} \left(\theta - \theta^{(j)}\right) + \dots + \frac{\mathrm{d}^M f\left(x_1, \theta\right)}{\mathrm{d}\theta^M} \bigg|_{\theta = \theta^{(j)}} \\ &\times \left(\theta - \theta^{(j)}\right)^M + \xi_1, \\ y_2 - f\left(x_2, \theta^{(j)}\right) \\ &= \frac{\mathrm{d}f\left(x_2, \theta\right)}{\mathrm{d}\theta} \bigg|_{\theta = \theta^{(j)}} \left(\theta - \theta^{(j)}\right) + \dots + \frac{\mathrm{d}^M f\left(x_2, \theta\right)}{\mathrm{d}\theta^M} \bigg|_{\theta = \theta^{(j)}} \\ &\times \left(\theta - \theta^{(j)}\right)^M + \xi_2, \end{aligned}$$

$$y_{L} - f\left(x_{L}, \theta^{(j)}\right)$$

$$= \frac{\mathrm{d}f\left(x_{L}, \theta\right)}{\mathrm{d}\theta} \bigg|_{\theta = \theta^{(j)}} \left(\theta - \theta^{(j)}\right) + \dots + \frac{\mathrm{d}^{M} f(x_{L}, \theta)}{\mathrm{d}\theta^{M}} \bigg|_{\theta = \theta^{(j)}}$$

$$\times \left(\theta - \theta^{(j)}\right)^{M} + \xi_{L},$$
(26)

其中 ξ 是误差,L是样本系列长度.参数估计中,当第j步参数估计为 $\theta^{(j)}$,对新的估计参数 $\theta^{(j+1)}$ 使(26)式的误差平方和达最小是所希望的.即

$$\min_{\theta} \left\{ F = \sum_{i=1}^{L} \xi_i^2 \right\}. \tag{27}$$

$$\text{id } d_{ik} = \frac{\partial^k f\left(x_i;\theta\right)}{\partial \theta^k} \Bigg|_{\theta = \theta^{(j)}} \text{ All } \beta_k^{(j+1)} = \left(\theta^{(j+1)} - \theta^{(j)}\right)^k \,,$$

把 $\beta_{t}^{(j+1)}$ 作为新变量,有向量矩阵关系:

$$YF^{(j)} = DF^{(j)}\beta^{(j+1)} + \xi^{(j)},$$
 (28)

其中

$$\boldsymbol{\beta}^{(j+1)} = \begin{bmatrix} \theta^{(j+1)} - \theta^{(j)} \\ (\theta^{(j+1)} - \theta^{(j)})^2 \\ \vdots \\ (\theta^{(j+1)} - \theta^{(j)})^M \end{bmatrix}, \ \boldsymbol{YF}^{(j)} = \begin{bmatrix} y_1 - f(x_1; \theta^{(j)}) \\ y_2 - f(x_2; \theta^{(j)}) \\ \vdots \\ y_L - f(x_L; \theta^{(j)}) \end{bmatrix},$$

$$\boldsymbol{\xi}^{(j)} = \begin{bmatrix} \xi_1^{(j)} \\ \xi_2^{(j)} \\ \vdots \\ \vdots \end{bmatrix},$$

$$\mathbf{D}\mathbf{F}^{(j)} = \begin{bmatrix} \frac{\mathrm{d}f\left(x_{1};\theta^{(j)}\right)}{\mathrm{d}\theta} \bigg|_{\theta=\theta(j)} & \frac{\mathrm{d}^{2}f\left(x_{1};\theta^{(j)}\right)}{\mathrm{d}\theta^{2}} \bigg|_{\theta=\theta^{(j)}} \\ & \cdots & \frac{\mathrm{d}^{M}f\left(x_{1};\theta^{(j)}\right)}{\mathrm{d}\theta^{M}} \bigg|_{\theta=\theta^{(j)}} \\ & \frac{\mathrm{d}^{2}f\left(x_{2};\theta^{(j)}\right)}{\mathrm{d}\theta^{2}} \bigg|_{\theta=\theta^{(j)}} \\ & \cdots & \frac{\mathrm{d}^{M}f\left(x_{2};\theta^{(j)}\right)}{\mathrm{d}\theta^{M}} \bigg|_{\theta=\theta^{(j)}} \\ & \cdots & \cdots & \cdots \\ & \frac{\mathrm{d}f\left(x_{L};\theta^{(j)}\right)}{\mathrm{d}\theta} \bigg|_{\theta=\theta^{(j)}} & \frac{\mathrm{d}^{2}f\left(x_{L};\theta^{(j)}\right)}{\mathrm{d}\theta^{2}} \bigg|_{\theta=\theta^{(j)}} \\ & \cdots & \frac{\mathrm{d}^{M}f\left(x_{L};\theta^{(j)}\right)}{\mathrm{d}\theta^{M}} \bigg|_{\theta=\theta^{(j)}} \\ & \cdots & \frac{\mathrm{d}^{M}f\left(x_{L};\theta^{(j)}\right)}{\mathrm{d}\theta^{M}} \bigg|_{\theta=\theta^{(j)}} \end{bmatrix}$$

则有新变量的最小二乘估计:

$$\boldsymbol{\beta}^{(j+1)} = \left(\boldsymbol{D}\boldsymbol{F}^{(j)\mathsf{T}}\boldsymbol{D}\boldsymbol{F}^{(j)}\right)^{-1}\boldsymbol{D}\boldsymbol{F}^{(j)\mathsf{T}}\boldsymbol{Y}\boldsymbol{F}^{(j)},\tag{29}$$

其中 $\theta^{(j+1)}$ 是j+1步估计参数.

递推估计式为

$$\theta^{(j+1)} = \theta^{(j)} + \beta^{(j+1)}. \tag{30}$$

对于多参数函数的泰勒级数展开(24)式,类似地有(25)式的 *M* 次方参数函数表达、(26)式的观测方程,得类似于(30)式的递推估计.

4.3 截痕相交参数估计方法

- 一般多参数函数或模型的截痕相交参数估计方法计算流程见图 3, 具体步骤如下:
 - 1) 给定参数初值 $\theta^{(0)}$;
- 2) 据已知的参数向量计算导数矩阵 DF 和函数 向量 $f^{(i)}$;
 - 3) 用(30)和(31)式确定新的参数向量 $\theta^{(j+1)}$;
- 4) 判断 $\theta^{(j+1)}$ 是否最优值,如果是寻优结束,否则转 2)继续循环.

截痕相交法由于泰勒级数截断引起的误差,寻 找其优值过程也是逐步迫近过程. 初始参数值随机 生成,采用函数收敛、参数收敛及最大迭代次数限制 作为迭代终止条件.

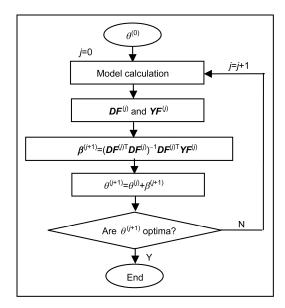


图 3 截痕相交参数估计方法计算流程

4.4 收敛性证明

截痕相交参数估计方法收敛性取决于(30)式中估计的新参数向量 $\theta^{(j+1)}$ 代入模型计算的值是否更接近于观测样本. 用误差平方和衡量其接近程度有

$$F^{(j+1)} = \xi^{(j+1)T} \xi^{(j+1)} = \left(y - f^{(j+1)} \right)^{T} \left(y - f^{(j+1)} \right), \quad (31)$$

其中 $\mathbf{y}^{\mathrm{T}} = [y_1 \quad y_2 \quad \cdots \quad y_L],$

$$\boldsymbol{f}^{(j+1)\mathrm{T}} = \begin{bmatrix} f(\boldsymbol{x}_1, \boldsymbol{\theta}^{(j+1)}) & f(\boldsymbol{x}_2, \boldsymbol{\theta}^{(j+1)}) & \cdots & f(\boldsymbol{x}_L, \boldsymbol{\theta}^{(j+1)}) \end{bmatrix}.$$

要证明(30)式新的寻找方向 $\Delta\theta = \theta^{(j+1)} - \theta^{(j)}$ 是正确的,只要证明(31)式对任意一步寻找的新参数向量 $\theta^{(j+1)}$ 相应的误差 $F^{(j+1)}$ 都比上一步误差 $F^{(j)}$ 更小即可.

证明如下.

$$F^{(j+1)} = \left(y - f^{(j)} - DF^{(j)}\beta^{(j+1)}\right)^{T} \left(y - f^{(j)} - DF^{(j)}\beta^{(j+1)}\right)$$

$$= F^{(j)} - (y - f^{(j)})^{T}DF^{(j)}\beta^{(j+1)} - \beta^{(j+1)T}DF^{(j)T}\left(y - f^{(j)}\right)$$

$$+ \beta^{(j+1)T}DF^{(j)T}DF^{(j)}\beta^{(j+1)}. \tag{32}$$

$$(30) 式代入(32) 式得$$

 $F^{(j+1)}$

$$= F^{(j)} - \beta^{(j+1)T} D F^{(j)T} (y - f^{(j)}) + \beta^{(j+1)T} D F^{(j)T} D F^{(j)} \beta^{(j+1)}$$

$$= F^{(j)} - (y - f^{(j)})^{T} D F^{(j)} (D F^{(j)T} D F^{(j)})^{-1} D F^{(j)T} (y - f^{(j)})$$

$$< F^{(j)}.$$

可知(29)和(30)式确定的任意一步新参数向量 $\theta^{(j+1)}$ 都满足关系:

$$F^{(j+1)} < F^{(j)}. (33)$$

由(33)式可知,每一步寻找,都会使误差平方和变得更小,随着寻找步骤增加,其误差平方和值越来越小,最终趋于最小值,获得最优的参数值.所以(30)式确定的寻找方向是正确的,其参数估计向量最终收敛于真参数.

5 截痕相交参数估计方法应用

根据以上分析与论证可知,如果模型不存在结构、参数与观测样本误差,那么总存在观测样本组使得样本组在模型参数域函数截痕间的交点是唯一的,而这交点就是模型真参数的位置.但实际模型存在诸多误差,这些误差会影响这截痕交点的位置,即影响估计的参数值.所以本研究对所提出方法检验分两步进行,第一步检验无误差理想函数参数估计效果,第二步论证有误差函数参数估计的有效性.

5.1 函数参数估计效果检验

为了方法检验简单又不失一般性,这里先对如下无误差函数的参数估计进行检验:

$$f(x;\theta) = 2 + (x\theta_1)^2 + x\theta_1 + \sin(2\pi x\theta_1) + (x\theta_2)^3 + \cos(2\pi x\theta_2).$$
(34)

该函数当自变量给定值 $x=x_i$ 时, 其样本函数是两维参数空间变化的函数:

$$f(x_i; \theta) = 2 + (x_i \theta_1)^2 + x_i \theta_1 + \sin(2\pi x_i \theta_1) + (x_i \theta_2)^3 + \cos(2\pi x_i \theta_2).$$
(35)

函数曲面分别随参数 \(\theta_i\) 的 2 次方、 \(\theta_i\) 的 3 次方和各自的正弦与余弦波动变化. 该函数的误差平方函数具有许多局部优值,其直接在目标函数曲面上寻找最优值十分困难.

为了论证方法估计参数的可行性, 检验其效果, 先对最简单的无误差函数进行验证. 假设参数值已知, 为 θ_1 =0.7, θ_2 =2.3;观测样本系列为: (x_1,y_1) , (x_2,y_2) ,···, (x_L,y_L) .

考虑到泰勒级数的阶数 M 会影响参数估计效果,这里分别采用 M=1,2,3,4,5 进行检验,当 M=1 即为高斯-牛顿方法.从这角度看,高斯-牛顿方法是截痕交点法阶数为 1 时的特例.

计算结果见表6和7.表6中j为循环次数,参数

下标的第二位表示泰勒级数的阶数. 表 7 中下标 G 表示用 Gauss-Newton 法获得的结果,下标 B 表示用截痕相交法获得的结果,J 表示参数估计循环次数.

从表 6 和 7 结果看,截痕相交法寻找新参数过程的每一步都十分有效,收敛速度很快,但采用泰勒级数的阶数越高,收敛速度越快.这是由于阶数越高的泰勒级数截断误差越小,收敛速度就越快,每步寻找新参数的效果越好.这说明以泰勒级数展开为基础的截痕交点法,不仅理论上比 Gauss-Newton 法更完善、适应面更广,应用效果上也更好.

5.2 函数参数估计方法效果比较

以上通过截痕交点法与 Gauss-Newton 法对无误差函数的参数估计进行了有效性检验,在理想情况下证明了所提出方法的可行性、有效性. 对于一种新的理论与方法, 还需要与传统的方法比较优缺点. 传统的确定性参数估计方法, 除与截痕交点法有类似理论基础的 Gauss-Newton 法外, 被公认为较好的方法有单纯形法(Simplex method)^[40]. 采用截痕交点法和单纯形法分别对函数(35)式进行参数估计,以分析比较不同方法的参数估计效果. 表 8 是两种方法相对于不同初始值开始的参数估计结果. 表 8 中下标 B 和 S 分别表示用截痕相交法和单纯形法获得的结果, J 表示参数估计从随机生成的初始值到找到其优值所循环的次数, 其他意义同前. 表 8 中同一行是用相同的参数初始值分别用截痕交点法和单纯形法求得优值,不同

表 6 一阶与 5 阶泰勒级数方法参数估计结果

J θ_{11} θ_{21} F_1 θ_{15} θ_{25} F_5 0 2.12 2.74 2.12 2.74 1 1.49 2.33 1248573 1.19 2.33 870266 2 1.21 2.29 3170 0.84 2.3 1007 3 1.05 2.29 1038 0.72 2.3 11 4 0.95 2.3 524 0.70 2.3 1.08×10 ⁻³ 5 0.89 2.3 321 321 321 321 321 321 321 322							
1 1.49 2.33 1248573 1.19 2.33 870266 2 1.21 2.29 3170 0.84 2.3 1007 3 1.05 2.29 1038 0.72 2.3 11 4 0.95 2.3 524 0.70 2.3 1.08×10 ⁻³ 5 0.89 2.3 321 6 0.84 2.3 205 7 0.82 2.3 183 8 0.79 2.3 159 9 0.77 2.3 114 10 0.74 2.3 38 11 0.70 2.3 0.614	J	θ_{11}	θ_{21}	F_1	θ_{15}	θ_{25}	F_5
2 1.21 2.29 3170 0.84 2.3 1007 3 1.05 2.29 1038 0.72 2.3 11 4 0.95 2.3 524 0.70 2.3 1.08×10 ⁻³ 5 0.89 2.3 321 6 0.84 2.3 205 7 0.82 2.3 183 8 0.79 2.3 159 9 0.77 2.3 114 10 0.74 2.3 38 11 0.70 2.3 0.614	0	2.12	2.74		2.12	2.74	
3 1.05 2.29 1038 0.72 2.3 11 4 0.95 2.3 524 0.70 2.3 1.08×10 ⁻³ 5 0.89 2.3 321 6 0.84 2.3 205 7 0.82 2.3 183 8 0.79 2.3 159 9 0.77 2.3 114 10 0.74 2.3 38 11 0.70 2.3 0.614	1	1.49	2.33	1248573	1.19	2.33	870266
4 0.95 2.3 524 0.70 2.3 1.08×10 ⁻³ 5 0.89 2.3 321 6 0.84 2.3 205 7 0.82 2.3 183 8 0.79 2.3 159 9 0.77 2.3 114 10 0.74 2.3 38 11 0.70 2.3 0.614	2	1.21	2.29	3170	0.84	2.3	1007
5 0.89 2.3 321 6 0.84 2.3 205 7 0.82 2.3 183 8 0.79 2.3 159 9 0.77 2.3 114 10 0.74 2.3 38 11 0.70 2.3 0.614	3	1.05	2.29	1038	0.72	2.3	11
6 0.84 2.3 205 7 0.82 2.3 183 8 0.79 2.3 159 9 0.77 2.3 114 10 0.74 2.3 38 11 0.70 2.3 0.614	4	0.95	2.3	524	0.70	2.3	1.08×10^{-3}
7 0.82 2.3 183 8 0.79 2.3 159 9 0.77 2.3 114 10 0.74 2.3 38 11 0.70 2.3 0.614	5	0.89	2.3	321			
8 0.79 2.3 159 9 0.77 2.3 114 10 0.74 2.3 38 11 0.70 2.3 0.614	6	0.84	2.3	205			
9 0.77 2.3 114 10 0.74 2.3 38 11 0.70 2.3 0.614	7	0.82	2.3	183			
10 0.74 2.3 38 11 0.70 2.3 0.614	8	0.79	2.3	159			
11 0.70 2.3 0.614	9	0.77	2.3	114			
	10	0.74	2.3	38			
$12 0.70 2.3 2.65 \times 10^{-4}$	11	0.70	2.3	0.614			
12 0.70 2.8 2.0010	12	0.70	2.3	2.65×10^{-4}			

表 7 Gauss-Newton 法与 5 阶截痕交点法结果比较

$\theta_{\scriptscriptstyle 1}{}^{\scriptscriptstyle (0)}$	$\theta_2^{(0)}$	$J_{ m G}$	$ heta_{ ext{IG}}$	$ heta_{ m 2G}$	$F_{ m G}$	$J_{ m B}$	$ heta_{ ext{1B}}$	$ heta_{ m 2B}$	$F_{ m B}$
0.99	0.984	4	0.7	2.3	0.008022	3	0.7002	2.3	3.83×10^{-3}
7.05	7.53	118	0.7019	2.3	6.97×10^{-2}	17	0.7001	2.3	2.10×10^{-4}
1.38	0.92	37	0.7003	2.3	3.83×10^{-3}	4	0.7003	2.3	3.66×10^{-3}
2.11	4	11	0.7001	2.3	3.26×10^{-4}	4	0.7002	2.3	5.20×10^{-3}
2.27	9.06	113	0.7019	2.3	7.48×10^{-2}	18	0.7001	2.3	2.17×10^{-4}
5.17	7.25	116	0.7028	2.3	0.15	16	0.7017	2.3	7.30×10^{-2}
1.36	0.52	55	0.7001	2.3	1.51×10^{-4}	3	0.7	2.3	0.004173
1.97	3.98	12	0.7	2.3	1.41×10^{-5}	4	0.7002	2.3	5.20×10^{-3}
2.5	0.24	68	0.7	2.3	1.16×10^{-5}	7	0.7	2.3	2.08×10^{-5}
4.08	9.92	115	0.7027	2.3	0.1407	17	0.7004	2.3	8.89×10^{-2}
1.38	3.74	6	0.7	2.3	2.30×10^{-5}	4	0.7	2.3	2.01×10^{-3}
1.17	1.37	14	0.7	2.3	4.57×10^{-5}	3	0.7003	2.3	3.89×10^{-3}
2.82	2.82	12	0.7001	2.3	2.69×10^{-4}	5	0.7001	2.3	1.30×10^{-4}
0.5	1.11	33	0.7	2.3	6.38×10^{-4}	3	0.7	2.3	1.81×10^{-3}
1.17	1.74	45	0.7	2.3	1.02×10^{-4}	3	0.7	2.3	6.00×10^{-4}
8.66	12.55	118	0.7028	2.3	0.1527	19	0.7008	2.3	2.99×10^{-2}
6.56	8.95	118	0.7017	2.3	0.0565	20	0.7019	2.3	8.03×10^{-2}
1.21	0.16	12	0.7	2.3	5.15×10^{-6}	7	0.7	2.3	2.49×10^{-5}
2.83	9.17	113	0.7017	2.3	0.0601	17	0.7001	2.3	1.97×10^{-4}
1.29	2.89	73	0.7	2.3	1.02×10^{-4}	4	0.7	2.3	4.61×10^{-6}
2.08	6.17	6	0.7001	2.3	8.02×10^{-4}	5	0.7	2.3	9.20×10^{-4}
1.28	0.95	14	0.7001	2.3	1.17×10^{-4}	4	0.7003	2.3	3.30×10^{-3}
2.76	3.4	12	0.7001	2.3	1.51×10^{-4}	5	0.7	2.3	3.37×10^{-6}
4.47	3.31	116	0.7017	2.3	0.0595	15	0.7024	2.3001	3.917
2.36	1.06	14	0.7	2.3	3.36×10^{-5}	5	0.7	2.3	2.24×10^{-5}
0.56	0.41	15	0.7	2.3	8.78×10^{-3}	3	0.7004	2.3	5.67×10^{-3}
4.23	4.95	115	0.7025	2.3	0.1282	15	0.7	2.3006	6.69×10^{-1}

行用随机生成的不同初始值;最后一行是平均的循 环次数和参数估计方差.

由表 8 可得两点规律性结论: 1) 截痕交点法的 平均循环次数(23.6 次)远比单纯形法(124.6 次)少. 说 明截痕交点法直接在函数曲面提取的参数估计信息 远比单纯形法在目标函数曲面提取的信息有效. 2) 截痕交点法估计的参数均方差(1.72×10 $^{-4}$, 1.01×10 $^{-4}$) 远小于单纯形法(1.08, 0.38), 证明截痕交点法不存在 局部优值, 而单纯形法明显存在局部优值. 表 9 列出了单纯形法的两个局部优值, F_1 , F_2 , F_3 和 F_4 是目标函数局部最小值 F_8 周围邻域的目标函数值. F_1 , F_2 , F_3 和 F_4 都比 F_8 大,说明确定的优值是其局部优值.

5.3 新安江模型应用检验

新安江模型的结构简单、物理基础强、应用于实际水流模拟效果好^[40-42],被广泛应用于流域的水流模拟与预报中.新安江模型参数优化率定得到了大量研究^[42-44].模型需要率定的灵敏参数有7个,见表10.以邵武流域为研究区域,位于中国福建省的闽江流域,气候湿润、台风暴雨与洪水频发.该流域1988~1999年的雨量站点密度较大、观测资料精度较高,且受人类活动影响小,所以选择该时期的水文日资料进行参数率定.参数率定结果见表11和12.表12列出了根据1000组随机生成的初始值优选参数获得的参数平均值和均方差,

表 8 不同方法参数估计效果比较分析

$J_{ m B}$	$ heta_{ ext{1B}}$	$ heta_{ ext{2B}}$	$F_{ m B}$	$J_{ m S}$	$ heta_{1 ext{S}}$	$ heta_{ m 2S}$	$F_{ m S}$
25	0.69994	2.29977	0.00314	109	1.031	2.292	402.7462
26	0.69993	2.30001	0.00963	150	0.881	2.295	156.24484
27	0.69993	2.29977	0.00276	52	0.701	2.299	0.04107
22	0.69994	2.3	0.00343	68	0.863	2.296	154.97693
25	0.69992	2.29977	0.00424	74	0.873	2.296	155.33882
24	0.70053	2.29975	0.00702	95	0.863	2.296	154.98294
25	0.69995	2.29977	0.00606	90	0.699	2.299	0.00077
17	0.70014	2.29977	0.00601	96	0.865	2.296	155.01048
20	0.70057	2.29975	0.00835	999	0.7002	2.299	0.003434
25	0.69993	2.29977	0.00359	84	0.857	2.296	154.80628
26	0.69993	2.3	0.00634	86	0.858	2.296	154.8277
29	0.70029	2.29976	0.00325	163	0.7	2.299	0.00529
26	0.69993	2.29977	0.00631	66	0.909	2.295	169.45652
27	0.69993	2.29977	0.00464	60	0.813	2.297	134.67823
20	0.70031	2.29976	0.00942	107	0.673	2.3	13.96529
26	0.69992	2.30001	0.0064	330	0.69993	2.3	0.008281
27	0.69993	2.29977	0.00304	50	0.892	2.295	158.85396
24	0.69996	2.29977	0.00733	58	0.866	2.296	155.03202
22	0.69993	2.29977	0.00672	46	1.041	2.291	416.61794
18	0.70036	2.29976	0.00373	96	0.842	2.296	152.94777
27	0.69993	2.29977	0.00709	80	0.699	2.299	0.00184
23	0.69993	2.29977	0.00563	91	0.7	2.299	0.01435
24	0.69992	2.29977	0.00435	77	0.694	2.3	0.61722
26	0.69993	2.29977	0.00878	335	0.704	2.299	0.42839
23	0.69998	2.30001	0.00639	65	0.864	2.296	155.0071
16	0.69998	2.29976	0.00519	116	0.892	2.295	159.05438
26	0.69992	2.30002	0.00654	12	0.675	2.3	11.61511
28	0.69993	2.29977	0.00372	49	0.761	2.298	56.72757
23	0.69996	2.3	0.00644	307	0.692	2.3	0.9744
26	0.69993	2.29977	0.00954	73	0.714	2.299	3.98913
26	0.69992	2.29977	0.00564	101	0.876	2.296	155.54838
19	0.69996	2.30001	0.00409	84	0.87	2.296	155.15329
23	0.69995	2.29977	0.00995	17	7.691	-0.164	2301122
17	0.69994	2.29976	0.00352	79	0.866	2.296	155.0425
27	0.69993	2.29977	0.00325	156	0.704	2.299	0.46433
16	0.70034	2.29976	0.00391	51	0.852	2.296	154.5392
21	0.70007	2.29977	0.00943	91	0.873	2.296	155.31394
18	0.69995	2.29977	0.0075	139	0.706	2.299	0.72201
25	0.69992	2.29977	0.00515	96	0.883	2.295	156.42645
27	0.69992	2.30001	0.0058	86	0.69	2.3	1.77654
23.6	1.72×10^{-4}	1.01×10^{-4}		124.6	1.08	0.38	

表 9 单纯形法局部优值

	θ_{1S} θ_{2S}	F_{S}	F_1	F_2	F_3	F_4
1.0	041 2.29	416.61794	416.7442	416.65793	416.71197	416.63334
7.0	691 -0.164	2301122.78	2301122.81	2301122.81	2301122.83	2301122.828

表 10 新安江模型参数

Parameter	Physical meaning
K	ratio of potential evapotranspiration to pan evaporation
B	exponent of watershed storage capacity curve
SM	areal mean free water storage capacity
CS	coefficient for surface flow concentration
CI	coefficient for interflow concentration
CG	coefficient for ground water concentration
KE	travel time through the reach (storage coefficient)

表 11 不同参数初值获得的参数优值

Parameter	Initial value (1)	Optimal value	Initial value (2)	Optimal value (2)
K	2.202199	1.111977	0.2908736	1.112445
B	0.1766915	1.013977	0.6608953	1.035315
SM	46.76068	29.06477	3.012237	28.90373
CS	0.1963299	0.3572581	0.292216	0.3577695
CI	0.799014	0.4521702	0.7813784	0.4543196
CG	0.9776084	0.9798456	0.9236649	0.9798568
KE	2.761958	1.459411	2.497829	1.458727
Average iteration number		9		40

表 12 参数方差统计

Parameter	Mean value of parameter	Mean square error	Relative mean square error (%)
K	1.11184	0.00299	0.269
B	1.01549	0.00984	0. 969
SM	29.14652	0.16194	0.556
CS	0.35701	0.00055	0.154
CI	0.45237	0.00234	0.517
CG	0.97977	0.00016	0.016
KE	1.45938	0.00108	0.074
Average iteration number	26.8		0.291

从表 11 和 12 计算结果可总结出: 1) 模型参数不存在局部优值; 2) 方法从不同的初始值获得的参数优值估计非常稳定,相对均方差(参数均方差与其均值之比)平均为 0.291%,最大的也只有 0.969%,最小的为 0.016%; 3) 寻找方法非常有效,平均寻找次数只需要 26.8,最多的也不超过 60 次,最小的循环次数只有 7 次循环就可以找到参数优值.

6 结语

通过大量具体实例分析了误差平方和目标函数和现有参数优化方法所存在的问题,发现:1)误差

平方和目标函数提供的参数估计信息存在诸多不合理性; 2) 对于变参数模型,目标函数会提供错误的参数估计信息; 3) 现有的参数优选方法都是在目标函数曲面上寻找,这由于误差平方和目标函数的结构使得问题复杂化,导致无法知道一般模型是存在唯一参数优值还是多优值,还会使唯一参数优值的模型在误差平方和目标函数曲面上变为多优值.误差平方和目标函数曲面上变为多优值.误差平方和目标函数曲面比参数函数自身曲面更复杂,参数估计不适合在目标函数曲面上寻找,而应在参数函数曲面上寻找.提出了参数函数、样本函数与样本截痕交点概念,分析论证了截痕交点唯一性定理(即参数优值的唯一性定理).

本文提出参数估计方法,比较现有方法具有如下特点和优点: 1) 直接在参数函数曲面上寻找参数.由于参数函数曲面比目标函数曲面结构更简单、方法更直接,使得参数估计过程更直观、获取的参数估计信息更可靠有效; 2) 用样本函数的泰勒级数获取的参数估计信息更精确.由于样本函数的泰勒级数获取的参数估计信息更精确.由于样本函数的泰勒级数反映了模型误差与参数间的精确关系,以该关系为基础获得的参数估计信息更精确、有效; 3) 方法的收敛性和参数估计的唯一性能严格地得到证明; 4)有效地解决或避免了现有参数估计方法的绝大多数问题.

当然,本文提出方法由于概念、理论基础、方法 是全新的,而且本文研究只以一个函数型参数和非 函数型模型估计为例,对于更多函数和模型的估计 效果、高阶导数估计方法等都有待深入研究,在实际 模型参数估计的应用也还需要做大量的适应性和有 效性检验.

附录 A

下面证明一般正整数指数函数的交点唯一性定理.

定理 A1. 任意 n 个参数,最高次为 m 的指数函数,最多需要 R(n,m)个样本,其截痕共同交点是唯一地相交于参数优值点.

其中

$$R(n,m) = \sum_{i=1}^{m} P(n,i)$$
, (A1)

其中 P(n,i) 是指数函数 $(\theta_1+\theta_2+\cdots+\theta_n)^i$ 展开的总项数, $P(n,i)=C_{i+n-1}^{n-1}$.

要证明定理 A1, 需要分别证明以下 3 个方面.

1) 任意 n 个参数 m 次方的指数函数, 最多有 R(n,m) 项参数项, 即证下式成立:

1060

$$R(n,m) \le \sum_{i=1}^{m} C_{i+n-1}^{n-1}$$
 (A2)

- 2) 任意 R(n,m)个样本截痕曲线能得参数唯一交点.
- 3) 截痕曲线唯一交点解等于参数优值.

证明 1: 采用数学归纳法证明, 当 m=1 时, 函数的一般表达式为

$$y(n,1) = a_1\theta_1 + a_2\theta_2 + \dots + a_n\theta_n + a_0$$
. (A3)

 $R(n,1) \leq C_n^{n-1} = n$ 成立;

当 m=2 时, 函数的一般表达式可表示为

$$y(n,2) = (a_1\theta_1 + a_2\theta_2 + \dots + a_n\theta_n)^2 + (a_1\theta_1 + a_2\theta_2 + \dots + a_n\theta_n)$$

$$+a_0$$
. (A4)

(A4)式可表示为

$$y(n,2) = (a_1\theta_1 + a_2\theta_2 + \dots + a_n\theta_n)^2 + y(n,1)$$
. (A5)

所以
$$R(n,2) \le P(n,2) + P(n,1)$$
 , 即 $R(n,2) \le \sum_{i=1}^{2} C_{i+n-1}^{n-1}$ 也

成立.

假设当 m=k 时,(A2)式仍然成立. 则当 m=k+1 时,函数的一般表达式可表示为

$$y(n,k+1) = (a_1\theta_1 + a_2\theta_2 + \dots + a_n\theta_n)^{k+1} + y(n,k),$$
 (A6)

$$R(n,k+1) \le C_{n+k}^{n-1} + R(n,k) \le C_{n+k}^{n-1} + \sum_{i=1}^{k} C_{n+k-1}^{n-1} = \sum_{i=1}^{k+1} C_{n+k}^{n-1}$$
 (A7)

因此, (A2)式得证.

证明 2: 对于任意 n 个参数 m 次方包含 R(n,m)项参数 项的指数函数,将 R(n,m)个样本代入函数式,就可得 R(n,m) 个方程组,此方程组把不同参数组合作为待求变量时属于线性系数方程组(例见文中(21)式),只要其系数矩阵满秩,就可直接求得参数组合变量解,由于参数组合变量中也包括了一次参数项,所以其一次项的解就是其唯一共同的交点,而此解在方程无误差时,显然就等于其真值.另外当系数矩阵不满秩时,即相同参数组合变量其系数组成的向

量间存在线性关系时,可以先消去线性相关的参数组合变量项(如文中(20)式中消去了 θ^2 与 β 两项,变成了新变量 θ^2 +2 β 和(21)式的满秩线性方程组)再求其解.所以,最多只需 R(n,m)个样本就可得唯一解,如果不满秩时,样本个数可小于 R(n,m)也可得唯一解.

证明 3: 由于每个样本是由真参数值代入指数函数计算得, 所以其解必然是最优参数值.

对于一般性函数或模型,下面进一步分析论证其截痕 交点的唯一性推理.

推理 A1: 对任意 n 维参数的函数,必存在 p 个不重叠的样本截痕组,使得样本组的截痕唯一地相交于参数优值点.

对于函数的不重叠样本截痕指的是在参数域上不处处相等的两个样本截痕. 对于一般性函数或模型,由于结构千差万别,直接进行严格的推导证明很难,这里采用分析推理方法论证. 对于任一观测样本的 y 样本函数 $y_i = f(x_i; \theta_1, \theta_2, \cdots, \theta_n)$ 是定义于n 维参数域上的,把样本函数描述的n 维空间区域记为 Y_i ,则对于任意 p 个不重叠截痕样本在 n 维空间区域的相交有如下关系:

$$\lim_{i \to n} \{ Y_1 \cap Y_2 \cap \dots \cap Y_i \} = \{ \theta \}, \tag{A8}$$

其中等号右边为截痕在参数优值的交点,等号左边表示 p 个样本截痕相交的公共部分. 由于对任意两个不重叠截痕样本,满足下式:

$$Y_i \cap Y_i < \min\{Y_i, Y_i\}. \tag{A9}$$

因此,随着相交样本截痕的增加,等式(A8)左边的公共部分迅速减少,当 *i=p* 时,其公共部分为一个点,即截痕在参数优值的交点。例如,一个3参数函数,其样本截痕为3维曲面,两个不重叠截痕样本相交的公共部分为曲线,3个相交的公共部分就只有有限的几个点,4个相交就为参数优值的交点。又如文中(10)式的任意两个不同的样本,其截痕在参数平面上最多只有4个交点,3个不同截痕的样本组就使其截痕的交点唯一地相交于参数优值点。

参考文献

- 1 Burke J V, Ferris M C. A Gauss-Newton method for convex composite optimization. Math Program, 1995, 71: 179–194
- 2 Rosenbrock H H. An automatic method of finding the greatest of least value of a Function. The Comp J, 1960, 3: 303-307
- Powell M J D. An efficient method for finding the minimum of a function of several variables without calculating derivatives. The Comp J, 1964, 7: 155–162
- 4 Pickup G. Testing the efficiency of algorithms and strategies for automatic calibration of rainfall-runoff models. Hydrol Sci Bull, 1977, 22: 257–274
- 5 Sorooshian S, Brazil E L. Comparison of Newton-type and direct search algorithms for calibration of conceptual rainfall-runoff models. Water Resour Res, 1988, 24: 691–700
- 6 Hofmann B. On the degree of ill-posedness for nonlinear problems. J Inverse and III-posed Prob, 1994, 2: 61-76
- 7 Hofmann B, Scherzer O. Factors influencing the ill-posedness of nonlinear problems. Inverse Probl, 1994, 10: 1277–1297
- 8 Beven K. A manifesto for the equifinality thesis. J Hydrol, 2006, 320: 18-36
- 9 Hooke R, Jeeves T A. Direct search solutions of numerical and statistical problems. J ACM, 1961, 8: 212-229
- 10 Ibbitt R P, Donnell T O'. Designing conceptual catchment models for automatic fitting methods. In: Proceedings of International

- Symposium on Mathematical Models in Hydrology, International Association of Hydrological Science. Warsaw, 1971
- 11 Francine M. Using a genetic algorithm combined with a local search method for the automatic calibration of conceptual rainfall–runoff models. Hydrolog Sci J, 1996, 41: 21–40
- 12 Cheng C T, Wang W C, Xu D M, et al. Optimizing hydropower reservoir operation using hybrid genetic algorithm and chaos. Water Resour Manag, 2008, 22: 895–909
- 13 Blum C. Ant colony optimization: Introduction and recent trends. Phys Life Rev, 2005, 2: 353-373
- 14 Yang J G, Zhuang Y B. An improved ant colony optimization algorithm for solving a complex combinatorial optimization problem. Appl Soft Comput, 2010, 10: 653–660
- 15 Chau K W. Particle swarm optimization training algorithm for ANNs in stage prediction of Shing Mun River. J Hydrol, 2006, 329: 363–367
- Jiang Y, Liu C M, Huang C C, et al. Improved particle swarm algorithm for hydrological parameter optimization. Appl Math Comput, 2010, 217: 3207–3215
- 17 Hopfield J, Tank D. 'Neural' computation of derisions in optimization problems. Biologieal Cyberneties, 1985, 52: 141-152
- 18 Duan Q Y, Sorooshian S, Gupta V. Effective and efficient global optimization for conceptual rainfall-runoff models. Water Resour Res, 1992, 28: 1015–1031
- 19 Duan Q Y, Gupta V K, Sorooshina S. Shuffled complex evolution approach for effective and efficient global minimization. J Optimiz Theory App, 1993, 76: 501–521
- 20 Gupta H V, Sorooshian S, Yapo P O. Status of automatic calibration for hydrologic models: comparison with multilevel expert calibration. J Hydrol Eng. 1999, 4: 135–143
- 21 Franchini M, Galeati G. Comparing several genetic algorithm schemes for the calibration of conceptual rainfall– runoff models. Hydrolog Sci J, 1997, 42: 357–379
- 22 Goswami M, O'Connor K M. Comparative assessment of six automatic optimization techniques for calibration of a conceptual rainfall-runoff model. Hydrolog Sci J, 2007, 52: 432–449
- 23 Xia Y, Pitman H V, Gupta M, Leplastrier A. Calibrating a land surface models of varyingcomplexity using multi-objective methods and the Cabauw data set. J Hydrometeorol, 2002, 3: 181–194
- 24 Engeland K, Braud I, Gottschalk L, et al. Multi-objective regional modeling. J Hydrol, 2006, 327, 339-351
- 25 Gupta H V, Sorooshian S, Yapo P O. Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information. Water Resou Res. 1998, 34: 751–763
- 26 包为民. 模型参数估计研究. 博士学位论文. 南京: 河海大学, 1989
- 27 包为民. 新安江模型参数的自动率定. 河海大学学报, 1986, 4: 22-30
- 28 Li Q, Bao W M, Zhang B, et al. Conceptual hierarchical calibration of Xinanjiang model. 3rd International Symposium on Methodology in Hydrology, IAHS Publ 350, 2011: 690–697
- 29 包为民. 具有有色噪声的模型参数估计. 水利学报, 1991, 12: 47-52
- 30 包为民, 稽海祥, 胡其美, 等. 抗差理论及在水文学中的应用. 水科学进展, 2003, 14: 133-137
- 31 包为民, 林跃, 黄贤庆, 等. 水库入库河段洪水汇流参数抗差估计研究. 武汉大学学报(工学版), 2004, 37: 13-17
- 32 赵超, 包为民, 王叶琴, 等. 河段汇流参数抗差估计研究. 河海大学学报, 2006, 34: 14-17
- 33 包为民, 王浩, 赵超, 等. AR 模型参数的抗差估计研究. 河海大学学报, 2006, 34: 258-261
- 34 Zhao C, Hong H S, Bao W M, et al. Robust recursive estimation of auto-regressive updating model parameters for real-time flood forecasting. J Hydrol, 2008, 349: 376–382
- 35 Bao W M, Li Q. Estimating selected parameters for the XAJ model under multicollinearity among watershed characteristics. J Hydrol Eng, 2012, 17: 118–128
- 36 包为民, 李倩, 李偲松. 基于灵敏度的条件目标函数构建方法研究. 水力发电学报(已接受)
- 37 李倩, 包为民. 基于傅里叶级数频率分解的汇流参数确定方法初探. 水电能源科学, 2011, 2: 33-38
- 38 包为民,司伟. 非线性参数的线性化率定方法. 计算力学学报, 2013, 32: 236-241 Bao W M, Li Q. An efficient calibration technique under the irregularity response surface. J Hydrol Eng, 2013, http://dx.doi.org/10.1061/(ASCE)AE.1943-5584.0000719
- 39 Nelde J A, Mead R. A simplex method for function minimization. The Computer J, 1965, 7: 308-313
- 40 Zhao R J. The Xinanjiang model applied in China. J Hydrol, 1992, 135: 371-381
- 41 Yao C, Li Z J, Bao H J. Application of a developed Grid-Xinanjiang model to Chinese watersheds for flood forecasting purpose. J Hydrol Eng, 2009, 14: 923–934
- 42 Hu C H, Guo S L, Xiong L H. A modified Xinanjiang model and its application in northern China. Nordic Hydrol, 2005, 36: 175-192
- 43 Cheng C T, Zhao M Y, Chau K W. Using genetic algorithm and TOPSIS for Xinanjiang model calibration with a single procedure. J Hydrol, 2006, 316: 129–140