第33卷第3期　　水 文　　Vol.33 No.3
2013年6月　　JOURNAL OF CHINA HYDROLOGY　　Jun., 2013

数据规律分析法在流量系数曲线公式拟合中的应用

卢玉成[1,2],王一帆[1,2],卢　曦[1,2]

(1.江苏省淮河入海水道管理处,江苏 淮安 223200;2.江苏省灌溉总渠管理处,江苏 淮安 223200)

摘　要:影响流量(效率)系数曲线的因素很多,给曲线配置公式,函数形式选择困难,待定系数计算复杂。通过对相关数据的统计分析,论证了数据规律与多项式函数对应关系的正确性,论证了数据规律与复合函数对应关系的正确性,论证了变量替换规则的正确性。运用这些关系和替换规则,引申出更多的函数形式,满足了一般曲线的公式配置要求。公式系数的求解,推导出降阶法,不但可以求解多项式类型函数的系数,还可以求解复合类型函数的系数;该方法统计计算简单,很容易用 Excel 表格实现。用数据规律分析的方法给相关曲线配置公式,非常有效。

关键词:水工建筑物;流量系数曲线;数据规律;多项式;复合函数;替换规则;降阶法;公式检验

中图分类号:O29;TV11　　**文献标识码**:A　　**文章编号**:1000-0852(2013)03-0069-08

1　概述

1.1　流量(效率)系数曲线公式拟合的意义

给水工建筑物流量(效率)系数曲线配置公式,可以直接转换成流量计算公式,这项工作有着积极意义。

(1)能够对曲线上节点的摘录误差进行一定的修正。正常从相关曲线上摘录节点,末位大部分为估读,存在一定的读数误差[1]。给曲线配置公式后,可以消除读数误差。

(2)可以作为推流曲线的延长方法应用。正常情况下相关曲线需要延长,延长的方法很多[2],而曲线配置的公式是该曲线段各种影响因素的综合反映,在一定区间内外延计算流量是合理的,较之趋势延长更加可靠。

(3)实现连续平滑推算流量。利用曲线的节点推算流量,应用的是插值法[3],例如流量整编程序应用的一元三点插值[4],其缺点是在临界点处会出现不连续或不平滑(左右导数不连续)现象[5],有些插值法[6]可以提高插值精度,但计算复杂,难以在推流程序中实现。利用配置公式计算流量不存在上述问题。

(4)极大提高报汛的时效性。水情信息的时效性要求很高[7],按照工作步骤,完成各水文要素观测,再到工作室利用关系表查算流量,编制报文,发送信息。如果观测点较远,时效很难保证。将推流曲线转变成公式,利用微型计算器的公式储存和计算功能,在观测点现场输入实测数据,就可以自动生成报文并及时发送,其水情的时效性将大大提高。在汛情紧张时刻,也可以为领导现场防汛决策提供实时计算成果。

(5)有利于相关曲线资料的储存、检索和应用。流量(效率)系数相关曲线图是水文资料整编的重要成果,以往只能保存手工绘制的纸质图纸,检索利用十分不便;曲线公式化以后,可以将公式填入流量率定成果表的附注栏,储存于水文数据库,便于检索和应用。必要时也可利用实测率定成果表的点据和配线公式由计算机按统一图幅格式成图,打印保存纸介质资料和存储成电子格式文档。

1.2　流量公式化方面的探索

建立流量与相关因素之间的函数公式,主要是公式形式选择和待定系数求解。众所周知,在曲线配置公式方面有较多的研究[8-10],模型函数的应用上也十分广泛,因水文要素影响因素复杂,但又不乏自身的特殊规律,这些方法能够成功地应用在推流关系中却很少。这方面,水文工作者也做了很多探索,如 Excel 试错法[11-12](用其添加趋势线功能进行公式选配)、多元对数逐步图解法[13]、组合函数分类迭代法[14]等。

1.3　本文曲线拟合方法——数据规律分析法

本文主要为经过单值化处理的水工建筑物流量

收稿日期:2012-04-16
作者简介:卢玉成(1956-),男,江苏沭阳人,高级工程师,从事水文测验整编方法研究。E-mail:596436303@qq.com

(效率)系数相关曲线选择适合的函数公式。以往此类二元函数拟合都是在已知公式类型的前提下用回归分析法[15],而曲线的公式形式很难确定,应用存在一定困难。本文首先分析数据规律与公式形式之间的关系,再经过变量替换,使公式形式包含更多的复合函数,极大地增加了公式的可塑性,能满足多种曲线形态的配型要求。待定系数计算,是配置公式的重要环节,本文通过对数据规律与数学公式之间关系的证明,推导出降阶法,只需将数据进行简单的统计计算即可求解待定系数。同时也解决了另一类型复合函数待定系数无法直接求解的问题。为应用方便,本文提供智能计算表格,只需将节点数据输入,可立即得到计算成果。本文用数据规律分析的方法给曲线配置公式是数理统计学中相关分析[16]内容的有益补充。

1.4 数据规律分析法的数据源

因实测资料的相关因素是随机的,没有一定的规律,因此流量(效率)系数相关曲线配置公式的数据源主要为曲线上按规定的方式摘录的节点,对于实测资料,按照规范要求[17]用以对公式的检验和误差评定。具体计算及检验方法,本文给出两类实例。

2 数据规律与数学公式

2.1 引理[1]

为研究数据规律与函数公式的关系,导入如下引理(证明略):

引理 对任何 $n,k\in N,a_1,a_2,\cdots,a_n\in R$,必有

$$\sum_{k=0}^{n}(-1)^{k}C_{n}^{k}\left(\sum_{i=0}^{n}a_{i}k^{i}\right)=(-1)^{n}a_{n}n! \qquad (1)$$

推论1 对任何 $n\in N,a,m\in R$,必有

$$\sum_{k=0}^{n}(-1)^{k}C_{n}^{k}(a-mk)^{n}=m^{n}n! \qquad (2)$$

推论2 对任何 $n\in N$,必有

$$\sum_{k=0}^{n-1}(-1)^{k}C_{n}^{k}(n-k)^{n}=n! \qquad (3)$$

2.2 数据规律与多项式的关系

在相关曲线上,我们只能摘取一部分节点,这些点是离散的。为了得到曲线的连续型函数表达式 $y=f(x)$,就必须对散点摘录作出规律性限制。在曲线段内,将 x_j 按等差(设差值为 δ)分配,$x_j=x_0+j\cdot\delta$,并摘录相应的 y_j。y_j 的各阶差值用 $M_{i,j}$ 表示,得出如下数据系列:

$$M_{i,j}=\begin{bmatrix} M_{1,1} & M_{1,2} & M_{1,3} & M_{1,4} & \cdots & M_{1,n} & \cdots & M_{1,m}\\ & M_{2,2} & M_{2,3} & M_{2,4} & \cdots & M_{2,n} & \cdots & M_{2,m}\\ & & M_{3,3} & M_{3,4} & \cdots & M_{3,n} & \cdots & M_{3,m}\\ & & & M_{4,4} & \cdots & M_{4,n} & \cdots & M_{4,m}\\ & & & & \ddots & \vdots & \cdots & \vdots\\ & & & & & M_{n,n} & \cdots & M_{n,m}\end{bmatrix} \qquad (4)$$

式中:$M_{1,j}=y_j-y_{j-1}$;$M_{i,j}=M_{i-1,j}-M_{i-1,j-1}$;$i$、$j$ 为自然数;$i\in[1,n]$,$j\in[i,m]$,$m\geqslant n$。可以看出:

$$M_{2,j}=M_{1,j}-M_{1,j-1}=(y_j-y_{j-1})-(y_{j-1}-y_{j-2})=y_j-2y_{j-1}+y_{j-2};$$

$$M_{3,j}=M_{2,j}-M_{2,j-1}=(y_j-2y_{j-1}+y_{j-2})-(y_{j-1}-2y_{j-2}+y_{j-3})=y_j-3y_{j-1}+3y_{j-2}-y_{j-3};$$

$$M_{4,j}=M_{3,j}-M_{3,j-1}=(y_j-3y_{j-1}+3y_{j-2}-y_{j-3})-(y_{j-1}-3y_{j-2}+3y_{j-3}-y_{j-4})=y_j-4y_{j-1}+6y_{j-2}-4y_{j-3}+y_{j-4};$$

…………

按照以上规律[2],利用逐步递推的办法,可以导出 $M_{i,j}$ 与 y_j 的关系,如下式:

$$M_{i,j}=\sum_{k=0}^{i}(-1)^{k}C_{i}^{k}y_{j-k} \qquad (5)$$

式(5)为滑动求和公式,i,j 属于定位参数,如当 $i=3$,$j=5$ 时,代入式(5)得:

$$M_{3,5}=\sum_{k=0}^{3}(-1)^{k}C_{3}^{k}y_{5-k}=y_5-3y_4+3y_3-y_2$$

结合上面定义,有如下定理:

定理1 对任何 $n,j\in N,a_0,a_1,a_2,\cdots,a_n,C,\delta,M_{n,j}\in R$,若 $M_{n,j}\equiv C\ (C\neq0)$,则函数必为如下 n 次方多项式:

$$y=a_nx^n+a_{n-1}x^{n-1}+\cdots\cdots+a_1x+a_0 \qquad (6)$$

且最高次项(首项)的系数为:$a_n=\dfrac{C}{n!\delta^{n}}$。

证 将各节点 $(x_0+j\delta,y_j)$ 代入式(6),得如下数据表:

$$\begin{bmatrix} y_0\\ y_1\\ \vdots\\ y_j\\ \vdots\\ y_m\end{bmatrix}=\begin{bmatrix} a_nx_0^n & +a_{n-1}x_0^{n-1} & +\cdots+a_1x_0 & +a_0\\ a_n(x_0+\delta)^n & +a_{n-1}(x_0+\delta)^{n-1} & +\cdots+a_1(x_0+\delta) & +a_0\\ \vdots & & & \\ a_n(x_0+j\delta)^n & +a_{n-1}(x_0+j\delta)^{n-1} & +\cdots+a_1(x_0+j\delta) & +a_0\\ \vdots & & & \\ a_n(x_0+m\delta)^n & +a_{n-1}(x_0+m\delta)^{n-1} & +\cdots+a_1(x_0+m\delta) & +a_0\end{bmatrix} \qquad (7)$$

在式(7)中,用差值法计算,左边为函数 y 的各阶差值,

---

[1] 见《数学中国论坛网》—基础数学 [1],作者:luyucheng,主题:求证一个 n! 的展开式。也可以直接点击如下网址查看具体证明过程:http://bbs.mathchina.com/cgi-bin/topic.cgi?forum=5&topic=14785。

[2] 各阶次用 y_j 表达式系数同二项式 $(a-b)^n$ 展开式的系数,且序号从 j 开始逐项降低。

右边差值采用同方次(或同系数)项差值。记 l 方次项的差值为 ${}^{l}M_{i,j}$,根据式(5)(需要将式中差值单元 y_{j-k} 替换成同指数项差值单元 $[a_l(x_0+(j-k)\delta)^l]$(k 为计算序列数)),第 l 方次项的第 l 次差值为:

$$
\begin{aligned}
{}^{l}M_{l,j}&=\sum_{k=0}^{l}(-1)^{k}C_l^k[a_l(x_0+(j-k)\delta)^l]\\
&=a_l\sum_{k=0}^{l}(-1)^{k}C_l^k((x_0+j\delta)-k\delta)^l
\end{aligned}
$$

令 $a=x_0+j\delta$,并根据引理中推论 1 的结论,有

$$
{}^{l}M_{l,j}=a_l\sum_{k=0}^{l}(-1)^{k}C_l^k(a-k\delta)^l=l!\delta^l a_l
$$

${}^{l}M_{l,j}=l!\delta^l a_l$ 是一个通项式,对 $l\in[1,n]$都成立。差值法,实际就是逐步消元法。式(7)中,一阶差值,消去 a_0 项,使x^1 项变为常数 $1!\delta^1a_1$;二阶差值,消去 x^1 项,使 x^2 项变为常数 $2!\delta^2a_2$;依此类推,可以逐步消元。当多项式为 n 次多项式时,n 阶差值消去 x^{n-1} 以下各元,多项式右边只剩下 n 次方的 n 阶差值 $n!\delta^n a_n$,多项式左边为 y 的 n 阶差值 C,因此,$a_n=\dfrac{C}{n!\delta^n}$,从而证明了首项系数公式的正确。

以上是按照多项式推导出定理 1 中的首项系数,由此也证明了:x 成等差变化,且 y 的 n 阶差值为固定常数时,函数为 n 次多项式。

2.3　数据规律与复合函数的关系

当 x 成等差变化时,y 的 u 阶差值的比值成常数 D,数据关系所对应的公式形式为指数函数和多项式函数组成的复合函数。为表述方便,记 ${}^{e}M_{i,j}$ 为 Be^{Ax_j}的 i 阶差值,记 ${}^{e'}M_{i,j}$ 为 e^{Ax_j}的 i 阶差值。有如下定理:

定理2　对任何$u,j\in N,A,B,a_0,a_1,a_2,\cdots,a_{u-1},D,\delta,M_{u,j}\in R$,若$\dfrac{M_{u,j+1}}{M_{u,j}}\equiv D(D>0,D\neq1)$,则函数必为如下复合函数:

$$
y=Be^{Ax}+a_{u-1}x^{u-1}+\cdots\cdots+a_1x+a_0 \tag{8}
$$

且 $A=\dfrac{\ln D}{\delta}$,$B=\dfrac{M_{u,j}}{{}^{e'}M_{u,j}}$。

证:从定理 1 的证明知道,差值法就是消元法。对于式(8)中多项式部分,u 阶差值,消去 x^{u-1} 以下各元,其左右两边的差值(右边仍按式(5)计算,须将差值单元 y_{j-k} 替换成指数项差值单元 $Be^{A(x_0+(j-k)\delta)}$)为:

$$
\begin{aligned}
M_{u,j}&=\sum_{k=0}^{u}(-1)^{k}C_u^k(Be^{A(x_0+(j-k)\delta)})\\
&=B\sum_{k=0}^{u}(-1)^{k}C_u^k e^{A(x_0+(j-k)\delta)}
\end{aligned}
$$

$$
\begin{aligned}
M_{u,j+1}&=\sum_{k=0}^{u}(-1)^{k}C_u^k(Be^{A(x_0+((j+1)-k)\delta)})\\
&=B\sum_{k=0}^{k=u}(-1)^{k}C_u^k e^{A(x_0+(j-k)\delta+\delta)}\\
&=Be^{A\delta}\sum_{k=0}^{u}(-1)^{k}C_u^k e^{A(x_0+(j-k)\delta)}
\end{aligned}
$$

故

$$
\frac{M_{u,j+1}}{M_{u,j}}=e^{A\delta}
$$

因为 y 的 u 阶差值的比值$\dfrac{M_{u,j+1}}{M_{u,j}}=D$,从而有 $D=e^{A\delta}$,所以 $A=\dfrac{\ln D}{\delta}$。当 A 确定后,代入式(8),u 阶差值后,左边为 $M_{u,j}$,右边为 ${}^{e}M_{u,j}=B\cdot{}^{e'}M_{u,j}$,所以,$B=\dfrac{M_{u,j}}{{}^{e'}M_{u,j}}$。

系数公式正确,也证明了数据规律与函数公式之间的对应关系成立。

2.4　替换规则与函数线型的扩展

定理 1 和定理 2 是研究常规下的数据规律与函数公式之间的关系,当数据进行一定的处理后,有如下定理:

定理 3(替换规则)

如果 x 按等差变化,y 取对数后满足定理 1 和定理 2 的规律,只需将函数式中的 y 用 $\ln(y)$替换即可;

如果 x 按等比变化(等比值为 q,系数计算中 $\delta=\ln(q)$),y 满足定理 1 和定理 2 的规律,只需将函数式中的 x 用 $\ln(x)$替换即可;

如果 x 按等比变化,y 取对数后满足定理 1 和定理 2 的规律,需将函数式中的 x、y 分别用 $\ln(x)$和 $\ln(y)$进行替换。

显而易见,替换后,按照定理 1 和定理 2 的证明方法,同样证明替换规则的正确性,此处不再重复。

定理 1 和定理 2 给出常规情况下的数据规律与函数公式之间的对应关系,而定理 3 则扩展了定理 1 和定理 2 的适用范围,这样可以变换出更多较为复杂的综合函数。

2.5　几个常见函数

定理 1、定理 2、定理 3 的联合使用,可以组合多种函数式,当然也包含常见函数,下面根据数据规律推导整理出几个常见函数。

当 x 成等差变化,对应的 y 成等比变化(即 y 取对数后一阶差为常数),则有:

$$
\begin{aligned}
\ln(y)&=+a_1x+a_0\\
y&=(e^{a_0})e^{a_1x}
\end{aligned}
$$

此为标准的指数函数。

当 x 成等比变化时，若函数 y 成等差变化(即 y 一阶差值为常数)，则有：

$$y=a_1\ln(x)+a_0$$

此为对数函数。

当 x 成等比变化，对应的 y 也成等比变化(即 y 取对数后一阶差值为常数)，则有：

$$\ln(y)=a_1\ln(x)+a_0$$

$$y=(e^{a_0})x^{a_1}$$

此为幂函数。

事实上，它们只是定理 1 中一次函数经过定理 3 中的几种替换得到的，此结果与 Excel 中提供的公式完全一致。

3 应用方法

前面已经从理论上证明了数据规律和函数形式的对应关系，但要实现公式化，对待定系数还需要选择一定的方法进行求解。实际配置过程中，为减少计算量和得到较好的结果，需作出相关要求。

3.1 待定系数计算(降阶法)

符合定理 1 的情况：本文推导数据规律与数学公式之间的关系，也直接推导出首项系数(n 方次)的计算公式，对 $n-1$ 方次项系数的求解，须将首项系数代入式(6)并将该项移至左边，使式(6)成为 $n-1$ 次多项式，再用 $n-1$ 阶差值即可计算，以下各方次系数计算按顺序类推。这种方法必须是从高次到低次逐个计算，称之为“降阶法”。分项系数计算见式(9)。

$$\left.\begin{aligned}
&a_n=\frac{C}{n!\delta^n}\\
&a_{n-1}=\frac{M_{n-1,j}-{}^{n}M_{n-1,j}}{(n-1)!\delta^{n-2}}\\
&a_{n-2}=\frac{M_{n-2,j}-{}^{n}M_{n-2,j}-{}^{n-1}M_{n-2,j}}{(n-2)!\delta^{n-2}}\\
&\quad\vdots\\
&a_p=\frac{M_{p,j}-{}^{n}M_{p,j}-{}^{n-1}M_{p,j}-\cdots-{}^{p+1}M_{p,j}}{p!\delta^p}\\
&\quad\vdots\\
&a_1=\frac{M_{1,j}-{}^{n}M_{1,j}-{}^{n-1}M_{1,j}-\cdots-{}^{2}M_{1,j}}{\delta}\\
&a_0=y_j-a_nx_j^n-a_{n-1}x_j^{n-1}-\cdots-a_1x_j
\end{aligned}\right\}\tag{9}$$

符合定理 2 的情况：此类函数的待定系数不能直接求解，以往多用迭代法反复计算，最终收敛于固定值。本文的方法十分简单，先求系数 A、B，然后用式(9)的方法计算其它各项系数。该法也是按顺序计算各个系数，应属“降阶法”范畴。具体计算公式见式(10)。

$$\left.\begin{aligned}
&A=\frac{\ln D}{\delta}\\
&B=\frac{M_{u,j}}{{}^{e'}M_{u,j}}\\
&a_{u-1}=\frac{M_{u-1,j}-{}^{e}M_{u-1,j}}{(u-1)!\delta^{u-1}}\\
&a_{u-2}=\frac{M_{u-2,j}-{}^{e}M_{u-2,j}-{}^{u-1}M_{u-2,j}}{(u-2)!\delta^{u-2}}\\
&\quad\vdots\\
&a_p=\frac{M_{p,j}-{}^{e}M_{p,j}-{}^{u-1}M_{p,j}-\cdots-{}^{p+1}M_{p,j}}{p!\delta^p}\\
&\quad\vdots\\
&a_1=\frac{M_{1,j}-{}^{e}M_{1,j}-{}^{u-1}M_{1,j}-\cdots-{}^{2}M_{1,j}}{\delta}\\
&a_0=y_j-Be^{Ax_j}-a_{u-1}x_j^{u-1}-\cdots-a_1x_j
\end{aligned}\right\}\tag{10}$$

经过定理 3 替换的情况：需要将要替换变量的数据系列进行处理后，仍然按式(9)、(10)进行系数求解。

3.2 比较阶次选择和摘录点据要求

公式阶次的选择：要考虑流量(效率)系数曲线是单弧线的特点和数据规律的分析比较不宜太复杂，公式的待定系数不宜太多，计算要相对容易。根据经验，定理 1 和定理 2 中的差值阶次不宜超过 3 阶。以三阶统计，就有 24 个公式形式和 84 个调节参数，是能够满足一般配线要求。

摘录点据的要求：阶次分析时最终比较的数据序列数不少于 2 个，按 3 阶比值情况统计，摘点总数应在 6 个以上(含首尾 2 点)，并且能控制曲线的变化规律。我们知道，流量(效率)系数曲线中下部曲率较大，上部曲率很小且变化平缓，正常情况下，自变量按等差摘录点据能控制曲线形态，在转换成同点数自变量按等比分布的点据一般也能控制曲线形态，如果曲线形态不符合这个规律，自变量按等比分配的点据应重新选择摘录。如果整线段数据分析没有得到好的结果，可从各阶差数据变化情况入手，进行分段处理，分段处理时每段曲线摘录点数也应不少于 6 个，且相邻曲线段分别至少有一个重合节点，确保曲线过度光滑连续。

3.3 数据计算中注意事项

在理论上，定理 1 中 n 阶差值 C 和定理 2 中 u 阶

差值的比值 D 是相同的一组常数。实际上它们只是一组近似数值,应该用该组数据的平均值,即 $C\approx\frac{1}{m-n}\sum_{j=n}^{m}C_j,D\approx\frac{1}{m-u-1}\sum_{j=u+1}^{m}D_j$,后面其它待定系数用同样方法计算,在统计学上的意义就是取系数的算术平均值。

在 x 成等比变化时,为减少计算带来的误差(因有效数字位数取舍不同,取对数后各差值略有差别),δ 取多点平均值,即 $\delta=(\ln x_m-\ln x_0)/m$,这样计算出的结果更具代表性。

4 配线公式检验

4.1 公式检验

在给相关曲线选配公式时,会不会出现局部反曲,用多个公式表达时,两段间过渡是否连续平滑,就需要对所选配的公式进行检验。我们设两函数的交点为(a,b),左段函数为 $f_1(x)$,右段函数为 $f_2(x)$。检验的充要条件为:

① 连续性检验:在相交点处,须 $f_1(a)=f_2(a)$。正常情况下流量取 3 位有效数字,计算值也可取三位有效数字进行比较。

② 平滑度检验:在相交点处满足 $f_1'(a)\approx f_2'(a)$。可取至一位小数进行比较,关键是它的变化规律要相衔接。

③ 无反曲检验,在线段上,$y''\neq 0$(y''恒大于 0 或恒小于 0)。

如果曲线只配有一个公式,只做③检验。

4.2 流量整编中检验

流量定线时,通过点群中心勾勒出相关曲线,摘取对应的点据,按照数据规律配置的经验公式,除做 4.1 的检验,还应对公式进行流量定线的三种检验,如果不合格,则应重新选配。

4.3 误差计算

新定线所配的公式,按规范要求统计实测点据的平均误差、最大误差和标准差;对已经使用的相关曲线配置的公式,要进行计算值与摘录值之间相对误差计算,这个误差可能含有摘录值的读数误差,因摘录值作为推流计算的真值,仍以其为评判标准,建议平均误差 $\left|\overline{X}\right|\leqslant 0.5\%$,最大误差 $\left|X_{\max}\right|\leqslant 1.0\%$,标准差 $S_e\leqslant 1.0\%$。

5 相关曲线公式配置优选算例

5.1 曲线公式的计算和选比

为了减少数据规律的分析过程,并且能够由曲线上节点迅即得到满意的配线成果,已经用 Excel 制成智能表格,应用时将曲线上节点输入(或粘贴)到表中指定位置,可立即得到 24 个公式的系数和误差评定值。选择误差评定最好的一个表达式为曲线的配置公式,也可以选择误差较小,表达式简单的作为配线公式。

智能表格分 6 个区块。数据源部分分成 2 个区块,应用时 x 等差分配栏 1、2 列需要输入(或粘贴)节点数据,第 3 列由第 2 列计算而得;x 等比分配栏可以重新输入节点数值,本文是用上栏数值按同节点数计算而得(内插点用一元三点插值法),同样第 3 列由 2 列计算而得。成果计算部分分成 4 个大的区块(共占 118 列),这 4 个区块除调用基础数据不同外,其它单元格计算方式完全一致;这 4 个区块中每区按阶差法和阶差比值法分别生成 6 个成果分表。表中各栏目(单元格)计算按照表头要求即可;需要说明的是,待定系数(最大误差)项,阶差法中非误差栏内,用上一格的数值除以相应的阶次系数 $n!\cdot\delta^n$(n 为对应的阶次),在误差栏,此单元格为最大误差;阶差比值法非误差栏内,A 为上一格的数值取对数后除以 δ,B 直接用上一格数值,后面的系数计算同阶差法,在误差栏内,此单元格为最大误差。具体布设见表 1。

若将本文的方法编制成计算机程序计算,十分简单。①创建节点总数和对应节点数值的输入界面;②将表 1 的数据源与核心区块设置成一个数组,用嵌套循环语句和函数计算功能将数组各单元格赋值;③依据误差评定栏的数据提示在指定误差范围内的成果供选择;④按选择结果编制成果表输出到打印机或输出到独立的文件。

5.2 析出成果示例及其检验

实例 1:淮安大引江闸(平板门平底闸)淹没式孔流线。相关因素为闸门开启高度 e 用 x 表示,流量系数 $M_2=\frac{Q}{Be\sqrt{\Delta Z}}$ 用 y 表示。相关线分为两段,x 在[0.2,1.2]内为曲线段,需选配合适的公式 $y=f_1(x)$;$x>1.20$ 时,流量系数稳定在 3.50,即 $f_2(x)=3.50$。

将节点输入表 1,比较得出 x 为等差、y 没有取对数、阶差法中的三阶 的结果为最优。从表 1 数据源中析出节点部分,从同区块一阶中析出 $M_{1,j}$、二阶中析出 $M_{2,j}$ 和三阶结果组合成表 2。

从表 2 待定系数(最大误差)栏知道,$a_3=-0.2083$,$a_2=1.7499$,$a_1=-3.4916$,$a_0=5.530$。

所以曲线的拟合公式是:$y=-0.2083x^3+1.7499x^2-$

表 1 数据配置公式计算(总体构成说明)

Table 1 The data configuration formula calculation (the structure description)

	数据源			y 没有取对数						y 取对数					
				阶差法(定理 1)			阶差比值法(定理 2)			阶差法(定理 1)			阶差比值法(定理 2)		
列序号	1	2	3	4–8	9–16	17–28	29–36	37–47	48–62	63–67	68–75	76–87	88–95	96–106	107–121
分表所占列数				(5 列)	(8 列)	(12 列)	(8 列)	(11 列)	(15 列)	(5 列)	(8 列)	(12 列)	(8 列)	(11 列)	(15 列)
数据代号	x_j	y_j	$\ln y_j$	一阶	二阶	三阶	一阶比	二阶比	三阶比	一阶	二阶	三阶	一阶比	二阶比	三阶比
数据源区与计算数据区	X 等差分配			单元格计算:原始数据取自前 1、2 列,其它数据取自本区域内已有的计算数据。应用定理 1 的公式形式。			单元格计算:原始数据取自前 1、2 列,其它数据取自本区域内已有的计算数据。应用定理 2 的公式形式。			单元格计算:原始数据取自前 1、3 列,其它数据取自本区域内已有的计算数据。应用定理 1 的公式形式和 $\ln(y)$替换。			单元格计算:原始数据取自前 1、3 列,其它数据取自本区域内已有的计算数据。应用定理 2 的公式形式和 $\ln(y)$替换。		
平均值(平均误差) 待定系数(最大误差) 标准差				成 果 统 计 计 算 区											
数据源区与计算数据区	X 为等比分配			单元格计算:原始数据取自前 1、2 列,其它数据取自本区域内已有的计算数据。应用定理 1 的公式形式和 $\ln(x)$替换。			单元格计算:原始数据取自前 1、2 列,其它数据取自本区域内已有的计算数据。应用定理 2 的公式形式和 $\ln(x)$替换。			单元格计算:原始数据取自前 1、3 列,其它数据取自本区域内已有的计算数据。应用定理 1 的公式形式和 $\ln(x)$、$\ln(y)$同时替换。			单元格计算:原始数据取自前 1、3 列,其它数据取自本区域内已有的计算数据。应用定理 2 的公式形式和 $\ln(x)$、$\ln(y)$同时替换。		
平均值(平均误差) 待定系数(最大误差) 标准差				成 果 统 计 计 算 区											

表 2 淮安大引江闸(淹没式孔流)经验公式选配计算

Table 2 Empirical formula with matching calculation of the empirical formula and data of Huaian Dayinjiang sluice (submersed sluice flow)

x_j	y_j	$M_{1,j}$	$M_{2,j}$	$M_{3,j}$	$a_3x_j^3$	${}^3M_{1,j}$	${}^3M_{2,j}$	$M_{2,j}-{}^3M_{2,j}$	$a_2x_j^2$	${}^2M_{1,j}$	$M_{1,j}-{}^3M_{1,j}-{}^2M_{1,j}$	a_1x_j	$y_j-a_3x_j^3-a_2x_j^2-a_1x_j$	验算 计算 M_2	验算 误差/%
0.20	4.90				–0.0017				0.070			–0.698	5.530	4.900	0.0
0.40	4.40	–0.50			–0.0133	–0.0117			0.280	0.210	–0.6983	–1.397	5.530	4.400	0.0
0.60	4.02	–0.38	0.12		–0.0450	–0.0317	–0.0200	0.1400	0.630	0.350	–0.6983	–2.095	5.530	4.020	0.0
0.80	3.75	–0.27	0.11	–0.01	–0.1066	–0.0617	–0.0300	0.1400	1.120	0.490	–0.6983	–2.793	5.530	3.750	0.0
1.00	3.58	–0.17	0.10	–0.01	–0.2083	–0.1017	–0.0400	0.1400	1.750	0.630	–0.6983	–3.492	5.530	3.580	0.0
1.20	3.50	–0.08	0.09	–0.01	–0.3599	–0.1516	–0.0500	0.1400	2.520	0.770	–0.6983	–4.190	5.530	3.500	0.0
平均值(平均误差)				–0.01				0.1400			–0.6983		5.530		0.0
待定系数(最大误差)				–0.2083				1.7499			–3.4916		5.530		0.0
标准差															0.0

$3.4916x+5.530$。

公式检验:

基本数据:$f_1(1.2)=3.50$,$f_2(1.2)=3.50$,$y'=-0.6249x^2+3.500x-3.492$,$y''=-1.250x+3.500$ 。

x_j	0.2	0.40	0.60	0.80	1.00	1.20	>1.2
$f'(x_j)$	–2.8	–2.2	–1.6	–1.1	–0.6	–0.2	0
$f''(x_j)$	3.3	3.0	2.8	2.5	2.3	2.0	

连续性检验:$f_1(1.2)=f_2(1.2)$,两线在相交点处连续;

平滑度检验:在 x 从 0.2 逐步增加到 1.2 时,曲线的斜率也从–2.8 逐步趋近于 0,所以两线过渡基本平滑;

无反曲检验:从基本数据看出,在(0.2,1.2)上,$y''>0$,所配公式不反曲。

误差评定:从表 2 的误差栏知道,平均误差、最大误差、标准差均为 0.0%,说明公式拟合曲线最好。

实例 2:六垛南闸(弧形门实用堰闸)淹没式孔流线,相关因素为 $e/\Delta Z$ 用 x 表示,流量系数 $M_2=\dfrac{Q}{Be\sqrt{\Delta Z}}$,用 y 表示。相关线分为 2 段,在[0.2,9.5]为曲线段,需配置公式 $f_1(x)$;$x>9.5$,$f_2(x)=4.94$ 。

将节点输入表 1,比较得出 x 为等比、y 取对数、阶差比值法中的二阶 的结果为最优。从表 1 数据源 x 为等比分配栏中析出 x_j、y_j、$\ln y_j$ 节点部分,从同区块一阶差中析出 $M_{1,j}$、二阶差中析出 $M_{2,j}$ 和二阶差比值法结果组合成表 3。

从表 3 待定系数(最大误差)栏知道,$A=0.8045$,$B=-0.04229$,$a_1=0.2404$,$a_0=1.315$。所以曲线的拟合公式是:$\ln(y)=-0.04229e^{0.8045\ln(x)}+0.2404\ln(x)+1.315$。

整理得: $y=3.7255\cdot x^{0.2404}\cdot 0.9586^{x^{0.8045}}$

公式检验:

基本数据:$f_1(9.5)=4.94$,$f_2(9.5)=4.94$,

表 3　六垛南闸(淹没式孔流)经验公式选配计算

Table 3　Empirical formula with matching calculation of the empirical formula and data of Liuduo south sluice (submersed sluice flow)

x_j	y_j	$\ln y_j$	$M_{1,j}$	$M_{2,j}$	$M_{2,j}/M_{2,j-1}$	$e^{A\ln(x_j)}$	$^{e'}M_{2,j}$	$M_{2,j}/^{e'}M_{2,j}$	$Be^{A\ln(x_j)}$	$^{e}M_{1,j}$	$M_{1,j}-^{e}M_{1,j}$	$a_1\ln x_j$	$\ln y_j-Be^{A\ln(x_j)}-a_1\ln x_j$	验算 计算 M_2	验算 误差%
0.200	2.50	0.916				0.2739			−0.01158			−0.3869	1.3148	2.50	0.0
0.433	2.98	1.092	0.176			0.5099			−0.02156	−0.0100	0.1856	−0.2013	1.3148	2.98	0.0
0.937	3.52	1.258	0.167	−0.0091		0.9489	0.2032	−0.04477	−0.04013	−0.0186	0.1851	−0.0157	1.3143	3.52	0.0
2.028	4.10	1.411	0.153	−0.0140	1.541	1.7661	0.3781	−0.03706	−0.07468	−0.0346	0.1871	0.1700	1.3157	4.10	0.0
4.389	4.63	1.533	0.122	−0.0310	2.209	3.2871	0.7038	−0.04399	−0.13900	−0.0643	0.1859	0.3556	1.3159	4.63	0.0
9.500	4.94	1.597	0.065	−0.0568	1.834	6.1179	1.3098	−0.04334	−0.25870	−0.1197	0.1845	0.5413	1.3148	4.94	0.0
平均值(平均误差)					1.861			−0.04229			0.1856		1.3152		0.0
待定系数(最大误差)					0.8045			−0.04229			0.2404		1.3152		0.0
标准差															0.0

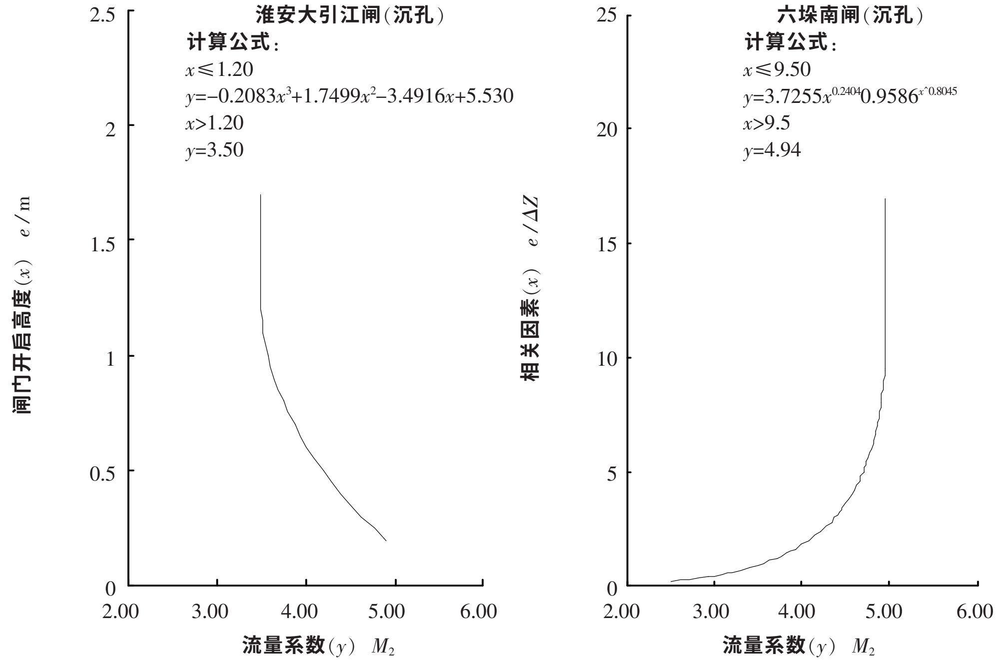

图 1　流量系数拟合曲线

Fig.1　Discharge coefficient fitting curve

$$y'=y(-0.0340x^{-0.1955}+0.2404/x),$$

$$y''=(y')^2/y+y(0.00665x^{-1.1955}-0.2404/x^2)。$$

x_j	0.200	0.433	0.937	2.028	4.389	9.500	>9.50
$f(x_j)$	2.50	2.98	3.52	4.10	4.63	4.94	4.94
$f'(x_j)$	2.9	1.5	0.8	0.4	0.1	0.02	0.0
$f''(x_j)$	−11.57	−2.98	−0.76	−0.20	−0.05	−0.01	

连续性检验：$f_1(9.5)=f_2(9.5)$，两线在相交点处连续；

平滑度检验：在 x 从 0.2 逐步增加到 9.5 时，曲线的斜率从 2.9 逐步趋近于 0，所以两线过渡平滑；

无反曲检验：从基本数据看出，在(0.2，9.5)上，$y''<0$，所配公式不反曲。

误差评定：从表 3 的误差栏知，平均误差、最大误差、标准差均为 0.0%，说明公式拟合曲线很好。

实例 1 和实例 2 的拟合与原图吻合。由于是利用函数表达式精细绘制，图形更加平滑美观。如需要读图，添加网格线后，可以随意缩放阅读，在指定的有效数字内几乎不存在读图误差。见图 1。

6　结语

利用本文的方法，对江苏省灌溉总渠管理处所属测站推流相关曲线共 18 条(不计直线部分，因为大部分流量系数曲线上部为稳定值)进行公式化处理，其类别有自由式孔流、淹没式孔流、淹没式堰流、水电站效率系数等相关曲线。二阶以下(含对数变换)的 14 条，说明大部分曲线数据结构简单，配线容易；用一个公式就能表达的 17 条，只有 1 条需用 2 个公式表达，即近

95%可配成单一线;符合定理 1 中多项式 3 条,常见函数 1 条,复合函数多达 14 条,可见大部分曲线是复合函数形式;所有配置公式拟合误差的平均值为 0.0 %,最大误差为 0.6%,标准差为 0.5%,说明用本文的方法拟合曲线的精度很高。

水文的方法也可用于其它相关曲线的公式配置。

参考文献:

[1] 林伟.水文勘测工[M].成都:电子科技大学出版社,2004:150.(LIN Wei. Hydrological Survey [M]. Chengdu: University of Electronic Science and Technology Press,2004:150.(in Chinese))

[2] 刘俊民.水文与水资源学[M].北京:中国林业出版社,2001:52-54.(LIU Junmin.Hydrology and Water Resources [M]. Beijing:China Forestry Press,2001:52-54.(in Chinese))

[3] 张光澄,张雷.实用数值分析[M].成都:四川大学出版社,2009:111-133.(ZHANG Guangcheng,ZHANG Lei. Practical Numerical Analysis [M]. Chengdu: Sichuan University Press,2009:111-133.(in Chinese))

[4] 赵玉然,岳树堂,刘克岩.水文要素过程整编程序化处理方法探讨与实践[J]. 河北工程技术高等专科学校学报, 2009,(2):1-4.(ZHAO Yuran, YUE Shutang,LIU Keyan.Discussion and practice of programmed method for process compilation of hydrological factors [J].Journal of Hebei Engineering and Technical College,2009,(2):1-4.(in Chinese))

[5] 常汶篇.一元三点插值法推流误差分析[J].水文,1982,7(1):23-28.(CHANG Wenpian.One yuan and three point interpolation method push flow error analysis [J].Journal of China Hydrology,1982,7(1):23-28.(in Chinese))

[6] 沈剑华.计算数学基础[M].上海:同济大学出版社,1989:41-60.(SHEN Jianhua.Foundations of Computational Mathematics [M].Shanghai: Tongji University Press,1989:41-60.(in Chinese))

[7] SL250—2000,水文情报预报规范[S].(SL250—2000,Hydrological Information and Forecast Norms[S].(in Chinese))

[8] 康来鹏. 实验数据处理中曲线拟合方法的探讨[J] . 真空电子技术,1992,181 (6):1-4,7.(KANG Laiping.Curve fitting in experimental data processing method discussion [J].Vacuum Electronics,1992,181(6):1-4,7.(in Chinese))

[9] 张继龙,甄蜀春,曹鹏,姚广锋.实验数据的曲线拟合方法及其应用[J].测试技术学报,2003,45(3):255-257.(ZHANG Jilong,ZHEN Shuchun, CAO Peng,YAO Guangfeng.The experimental data curve fitting method and its application [J].Journal of Test And Measurement Technology,2003,45(3):255-257.(in Chinese))

[10] 平书伟.关于曲线拟合的方法探讨[J] .机电信息,2011,294(12):112-113.(PING Shuwei.Study on curve fitting method[J].Mechanical and Electrical Information.2011,294(12):112-113.(in Chinese))

[11] 蔡亮.Excel 在单一线法定线中的应用[J].水文,2011,31(2):32-34,5.(CAI Liang.Excel in single line method of fixed line application [J].Journal of China Hydrology,2011,31(2):32-34,5.(in Chinese))

[12] 朱永军.EXCEL 软件在流量定线中的应用[J].江苏水利,2007,117(05):29-30.(ZHU Yongjun.EXCEL software in the application of flow line[J].Jiangsu Water Resources,2007,117(05):29-30.(in Chinese))

[13] 许嘉谟.多元对数图解法整编堰闸流量资料[J].水文,1981,6(6):21-25.(XU Jiamo.Multiple logarithm graphic method weir sluice flow data reorganization[J].Journal of China Hydrology,1981,6(6):21-25.(in Chinese))

[14] 卢玉成.回归分析在运西电站、节制闸流量整编中的应用[J].水文,1990,15(4):57-59.(LU Yucheng. Regression analysis in Yunxi power station and sluice flow data application [J].Journal of China Hydrology,1990,15(4):57-59.(in Chinese))

[15] 薛毅.数值分析[M].北京:北京工业大学出版社,2003:223-232.(XUE Yi.Numerical Analysis[M]. Beijing: Beijing University of Technology Press,2003:223-232.(in Chinese))

[16] 朱勇华.应用数理统计[M].武汉:武汉水利电力大学出版社,2000:151-214.(ZHU Yonghua.Application of Mathematical Statistics[M]. Wuhan:Wuhan Water Conservancy and Electric Power University Publishing, 2000: 151-214.(in Chinese))

[17] SL247—2000, 水文资料整编规范 [S].(SL247—2000,Code for Hydrologic Data Compilation[S].(in Chinese))

Application of Data Rule Analysis Method in Discharge Coefficient Curve Formula Process

LU Yucheng[1,2], WANG Yifan[1,2], LU Xi[1,2]

(1. *Huaihe River Mouth Management Division of Jiangsu Province, Huaian* 223200, *China*;
2. *Irrigation Canal Management Division of Jiangsu Province, Huaian* 223200, *China*)

Abstract: There are lot factors to effect discharge coefficient (or efficiency coefficient)curve . As for curve configuration in formula,it is difficult to select function form and the undetermined coefficient calculation is complex. Through statistical analysis of data ,this paper proved that the relationship between data rules and polynomial function is correct,the relation between data rules and compound function is correct,and variable substitution rules are correct. Using these relationships and replacing rules ,more functional form were developed to meet the general formula configuration requirements.

Key words: hydraulic structure; discharge coefficient curve; data law; polynomial; composition function; replacing rule; method of order reduction; formula inspection