

多输入/单输出线性函数模型 在河道流量演算中的应用初探

陈 科

(成都勘测设计研究院, 成都, 610072)

提 要 在河道流量演算中常常遇到多输入/单输出的情况, 常用的流量演算方法处理较为困难, 为此本文采用多输入/单输出线性函数模型进行河道流量演算, 每个输入的线性响应函数为切比雪夫多项式, 系统的记忆长度及线性响应函数多项式的展开阶数通过试错法确定, 各个输入与输出的互相关系数可为系统记忆长度的确定提供初解。本文用两输入/单输出线性函数模型对岷江上游干流姜射坝站至干流皂角湾站的河道流量进行演算, 并考虑支流杂谷脑河桑坪站的流量汇入, 在实际计算中同时采用自回归模型进行实时校正, 其计算结果是令人满意的。

关键词 多输入/单输出 河道流量演算 响应函数 系统记忆长度

在河道流量演算中常遇到多输入/单输出的情况, 处理较为棘手, 有效的方法之一是采用多输入/单输出线性函数模型(Multiple input-single output linear functional models, 简称 LFM)。

1 多输入/单输出线性函数模型

单输入/单输出函数模型的积分式可表示为:

$$y(t) = \int_0^{t_m} h(r) \cdot x(t-r) dr \quad (1)$$

式中 $y(t)$ ——输出函数;

$x(t)$ ——输入函数;

r ——滞时;

$h(r)$ ——响应函数, 当 $r < 0$, 则 $h(r) = 0$;

t_m ——系统记忆长度。

对线性时不变系统而言, 将上积分写为离散形式为:

$$y(t) = \sum_{r=0}^u h(r) \cdot x(t-r) \quad (2)$$

式中 u ——系统记忆长度的时段数。

当输出为几个相互独立的输入线性组合时, 就得到多输入/单输出线性函数模型, 其

公式为:

$$y(t) = \sum_{i=1}^L \sum_{r=0}^{u(i)} h^{(i)}(r) \cdot x^{(i)}(t-r) \quad (3)$$

式中 L ——输入函数的总数;

$x^{(i)}(t)$ ——第 i 个输入函数;

$h^{(i)}(r)$ ——第 i 个输入的响应函数;

$u(i)$ ——第 i 个输入的系统记忆长度。

模型的关键是响应函数, 它可以采用正交多项式来拟合, 并且在河道流量模拟演算中常用黑箱子法来求解响应函数。

设 $F = [P_n(t), n=0, 1, \dots, M]$ 为正交多项式, 则第 i 个输入的系统响应函数 $h^{(i)}(r)$ 可表示为:

$$h^{(i)}(r) = \sum_{n=0}^M a_n^{(i)} P_n(r) \quad (4)$$

式中 M ——多项式展开的阶数;

$a_n^{(i)}$ ——系数。

多输入/单输出线性函数模型的响应函数采用切比雪夫(Chebyshev)多项式, 其递推形式为:

$$P_{n+1}(t) = 2t \cdot P_n(t) - P_{n-1}(t) \quad t \in [-1, 1] \quad (5)$$

$$P_0(t) = 1$$

$$P_1(t) = t$$

$$P_2(t) = 2t^2 - 1$$

$$P_3(t) = 4t^3 - 3t$$

.....

就是说式(4)中的 $P_n(r)$ 为切比雪夫多项式(实际计算中需对 r 作归一化处理),将式(4)右边的式子代入式(3)经整理得:

$$y(t) = \sum_{i=1}^L \sum_{n=0}^M a_n^{(i)} \sum_{r=0}^{u(i)} P_n(r) x^{(i)}(t-r) \quad (6)$$

且令:

$$C_n^{(i)}(t) = \sum_{r=0}^{u(i)} P_n(r) x^{(i)}(t-r) \quad (7)$$

$$\text{则有 } y(t) = \sum_{i=1}^L \sum_{n=0}^M a_n^{(i)} C_n^{(i)}(t) \quad (8)$$

通过输入输出序列,采用最小二乘法,以实测

$$\begin{cases} \sum_{i=1}^L \sum_{n=0}^M a_n^{(i)} \sum_{t=u(i)}^N C_n^{(i)}(t) \cdot C_0^{(1)}(t) = \sum_{t=u(i)}^N Q(t) \cdot C_0^{(1)}(t) \\ \sum_{i=1}^L \sum_{n=0}^M a_n^{(i)} \sum_{t=u(i)}^N C_n^{(i)}(t) \cdot C_1^{(1)}(t) = \sum_{t=u(i)}^N Q(t) \cdot C_1^{(1)}(t) \\ \dots\dots\dots \\ \sum_{i=1}^L \sum_{n=0}^M a_n^{(i)} \sum_{t=u(i)}^N C_n^{(i)}(t) \cdot C_{11}^{(1)}(t) = \sum_{t=u(i)}^N Q(t) \cdot C_{11}^{(1)}(t) \\ \sum_{i=1}^L \sum_{n=0}^M a_n^{(i)} \sum_{t=u(i)}^N C_n^{(i)}(t) \cdot C_0^{(2)}(t) = \sum_{t=u(i)}^N Q(t) \cdot C_0^{(2)}(t) \\ \sum_{i=1}^L \sum_{n=0}^M a_n^{(i)} \sum_{t=u(i)}^N C_n^{(i)}(t) \cdot C_1^{(2)}(t) = \sum_{t=u(i)}^N Q(t) \cdot C_1^{(2)}(t) \\ \dots\dots\dots \\ \sum_{i=1}^L \sum_{n=0}^M a_n^{(i)} \sum_{t=u(i)}^N C_n^{(i)}(t) \cdot C_{11}^{(2)}(t) = \sum_{t=u(i)}^N Q(t) \cdot C_{11}^{(2)}(t) \end{cases}$$

从上面的方程组可以看出, $a_n^{(i)}$ 的系数矩阵是对称方阵,因此可以通过高斯-约当消去法对多项式的系数 $a_n^{(i)}$ 进行求解。多输入/单输出线性函数模型的系统记忆长度 $u(i)$ 及多项式展开阶数 M 的率定采用试错法进行,在这之前先计算各个输入与输出的互相关系数,可确定系统记忆长度的初值,通过不断迭代、试错可找到 $u(i)$ 和 M 的最优值。

实测计算中,实测值与预报值的残差序列为: $Z(t) = Q(t) - y(t)$, $Z(t)$ 应当为白噪声,当 $Z(t)$ 不为白噪声时,为了改进模拟模型,可采用自回归模型 $AR(P)$ 进行误差预报,以对预报值进行实时修正,即可用计算时

输出过程 $Q(t)$ 与计算过程 $y(t)$ 的平方和最小为目标函数:

$$E = \sum_{t=u(i)}^N \{Q(t) - [\sum_{i=1}^L \sum_{n=0}^M a_n^{(i)} C_n^{(i)}(t)]\}^2 = \min \quad (9)$$

式中 N ——使用的数据个数;

用式(9)对 $a_n^{(i)}$ 求偏导并令其为 0,即

$$\frac{\partial E}{\partial a_n^{(i)}} = -2 \sum_{t=u(i)}^N \{Q(t) - [\sum_{i=1}^L \sum_{n=0}^M a_n^{(i)} C_n^{(i)}(t)]\} \cdot C_n^{(i)}(t) = 0$$

可得到 $L \times (M+1)$ 个方程,例如,当 $L=2$ 、 $M=11$ 时,则共有 $2 \times (11+1) = 24$ (个)方程,如下:

段以前 P 个时段的误差作误差预报,并不断递推,对预报值进行实时修正。

这样,进一步地多输入/单输出线性函数模型式(3)及误差模型就构成了组合的多输入/单输出线性函数模拟模型(Combined model,简称 CLFM),对流域出口断面的输出流量进行计算,并可考虑流量计算的实时误差校正。

2 模型的有效性评定

模型的评定采用点绘实测值、计算值和校正后计算值过程线及分析计算统计参数

$RMSE$ 、 E_1 、 E_2 和 E_3 来进行。 $RMSE$ 是均方差, E_1 是误差平均值与观测流量平均值之比, E_2 误差平方的均值与观测流量平方的均值之比, E_3 观测值与计算值的相关系数, 它们的计算公式为:

$$RMSE = \left\{ \sum_{t=1}^N [y(t) - Q(t)]^2 / N \right\}^{\frac{1}{2}} \quad (10)$$

$$E_1 = \frac{1}{N} \sum_{t=1}^N Z(t) / \frac{1}{N} \sum_{t=1}^N Q(t) \quad (11)$$

$$E_2 = \frac{1}{N} \sum_{t=1}^N Z^2(t) / \frac{1}{N} \sum_{t=1}^N Q^2(t) \quad (12)$$

$$E_3 = \frac{\sum_{t=1}^N [Q(t) - \bar{Q}][y(t) - \bar{y}]}{\left(\sum_{t=1}^N [Q(t) - \bar{Q}]^2 \sum_{t=1}^N [y(t) - \bar{y}]^2 \right)^{-\frac{1}{2}}} \quad (13)$$

式中 $\bar{Q}$ 、 $\bar{y}$ ——实测值、计算值的均值。

当 E_1 、 E_2 值接近 0 时, E_3 接近 1 时, 模型成功。

3 应用举例

为了对本文介绍的多输入/单输出线性函数模型有进一步的认识, 以及了解它在实际流量演算中的适应性, 本文采用实际流域进行计算。

3.1 流域及使用资料

将多输入/单输出模型应用到岷江上游干流姜射坝站、支流杂谷脑河桑坪站和下游干流皂角湾站的流量模拟上, 该流域情况见图 1:



图 1 岷江干流流域图

姜射坝与桑坪至皂角湾的河长距离均约为 60km, 流域系统的两个输入分别是姜射坝站的流量过程和桑坪站的流量过程, 输出为皂角湾站的流量过程, 显然, 这符合两输入/单输出的情况。

本文的两输入/单输出模型的流量演算采用上述三个站的 1984、1985 年全年 365 天的日径流资料。

3.2 分析计算

3.2.1 响应函数的估计 本文的响应函数采用切比雪夫多项式, 响应函数的估计主要是对响应函数展开阶数 M 进行识别。在阶数识别时, 当 M 增加, 回归系数增加, 率定系统性改进。假设各个输入不同的记忆长度 $u(i)$ 及不同的 M 值, 计算 $RMSE$ 值, 在误差停止迅速减小处决定最优展开阶数。通过试错计算, 确定了本文两输入/单输出响应函数的展开阶数 $M = 11$ 。

3.2.2 系统记忆长度的确定 第 i 个输入的系统记忆长度 $u(i)$ 即第 i 个输入影响存在的时间间隔。一般情况下, 也是通过试错法确定, 在 M 一定的条件下, 不断改变 $u(i)$ 并同时计算 $RMSE$ 值, 在误差停止迅速减小处决定各个输入的 $u(i)$ 。

计算第 i 个输入和输出之间的互相关系数可为 $u(i)$ 的确定提供初解。互相关系数随滞时 k 的增加而递减, 本文的两输入/单输出的 X_1 、 X_2 (输入) 和 Y (输出) 的互相关关系见图 2, 在相关图上第一个最小正值附近决定总记忆长度, 在相关图上得到 $u(i)$ 值后, 再在其附近进行试算, 可以得到最优值, 其计

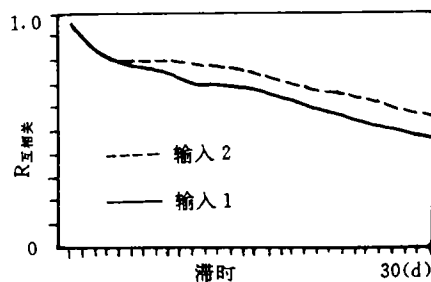


图 2 两个输入与输出的互相关关系图

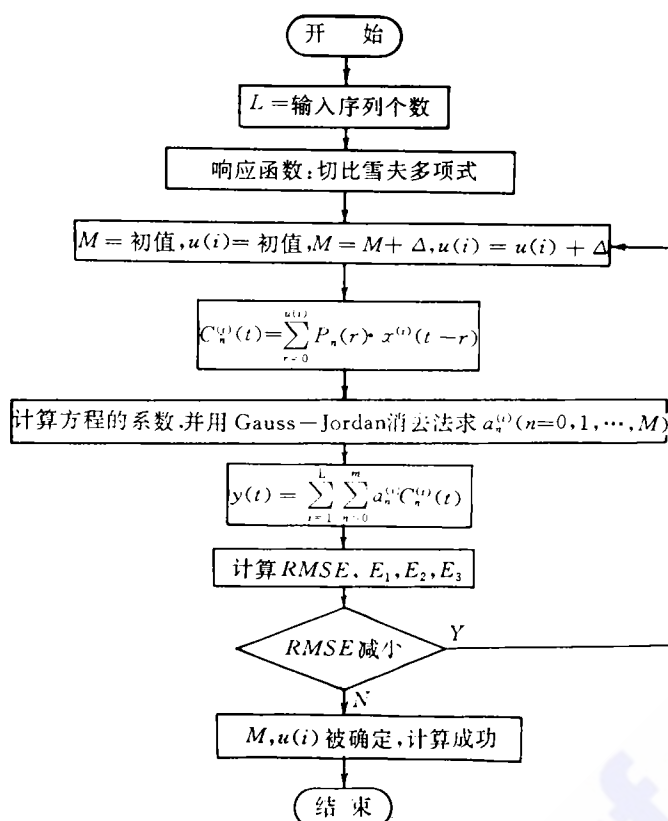


图3 多输入—单输出模型参数求定流程图

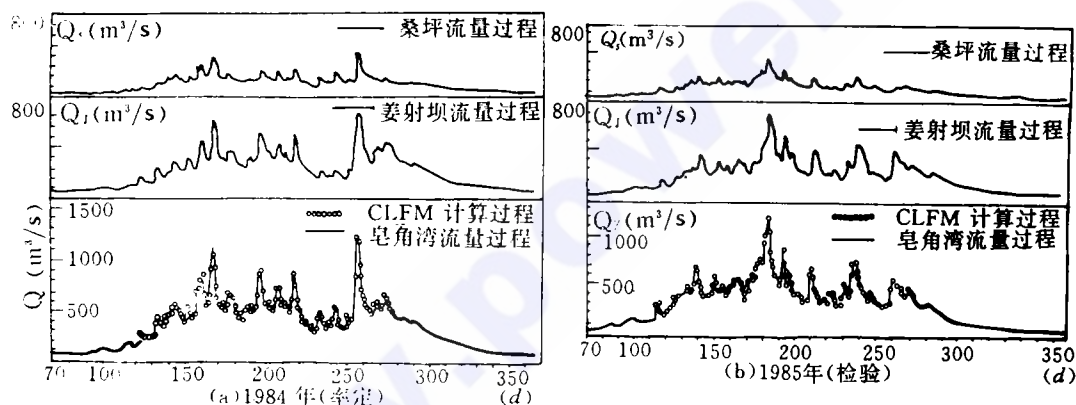


图4 实测与计算流量过程线图

算过程见图3。

通过计算,本文的两输入/单输出的两个输入的内存长度均为12(d),模型计算结果见图4、表1、2,从这些结果来看,其计算精度是很高的,模拟过程是令人满意的。

本文通过对实测过程与LFM模型计算

过程的残差序列的自相关分析、偏相关分析,确定了AR(P)模型的阶数 $P=2$ 即用2阶自回归模型进行误差预报,将其与LFM模型结合起来,构成组合模型(CLFM),其计算结果见图4、表1、2,可见模型的计算精度是很高的。

表 1 实测和计算洪峰流量比较(皂角湾站)

年份	序号	时 间 (月·日)	实 测 (m ³ /s)	LFM (m ³ /s)	CLFM (m ³ /s)
1984	1	3.29	208	200	208
	2	4.24	608	587	587
	3	5.10	888	883	884
	4	5.16	1090	1110	1074
	5	5.26	722	676	706
	6	6.16	905	911	926
	7	6.26	761	745	761
	8	7.5	928	887	888
	9	7.21	497	490	509
	10	7.30	548	538	549
	11	8.12	1250	1247	1248
	12	9.30	678	691	686
1985	1	3.10	217	195	205
	2	3.26	351	329	344
	3	4.5	406	379	386
	4	4.14	511	521	519
	5	4.20	725	686	691
	6	4.30	658	619	631
	7	5.15	600	574	575
	8	6.2	1240	1274	1243
	9	6.12	918	909	907
	10	6.30	730	718	730
	11	7.26	798	784	813
	12	8.18	667	638	634
	13	8.27	549	514	532

表 2 LFM 和 CLFM 模型评定成果表

年份	模 型	E_1	E_2	E_3	RMSE
1984	LFM	0.00	-0.00	1.00	13.13
	CLFM	0.00	-0.00	1.00	9.78
1985	LFM	0.00	-0.00	1.00	18.16
	CLEM	0.00	-0.00	1.00	15.28

4 几点认识

通过对多输入/单输出线性函数模型结构的分析研究和对岷江上游姜射坝、桑坪及皂角湾站的流量的实际计算,可以知道:

1. 多输入/单输出线性函数模型(LFM)用于河道流量演算是可行的;

2. 每个输入序列的线性响应函数可由切比雪夫多项式拟合,方程组的系数由观测数据估计;

3. 模型参数(如各个输入序列的系统记忆长度及线性响应函数的阶数)可通过试错法加以决定,每个输入与输出之间的互相关系数可为记忆长度的确定提供初解;

4. 两输入/单输出线性函数模型用于岷江上游干流姜射坝站、支流杂谷脑河桑坪站及下游干流皂角湾站的日流量演算,其结果是令人满意的;

5. 误差序列采用随机 AR 模型拟合处理,LFM 结合误差预报的 AR 模型进行计算,其精度比只用 LFM 计算的精度更高。

参 考 文 献

1 D. M. papamichail and Z. G. papazafirion. Multiple input
single output linear functional models for river flow
routing, Journal of Hydrology, Vol. 133, No. 3-4, May
1992

2 李庆杨、王能超、易大义编. 数值分析. 华中工学院出版社, 1986

3 丁晶、邓育仁编著. 随机水文学. 成都科技大学出版社, 1988

4 项静恬、杜金观、史久恩. 动态数据处理—时间序列分析. 气象出版社, 1986 年

(收稿日期 19930611)