

doi: 10.3969/j.issn.0490-6756.2020.05.014

# 基于 GMDH 的宏观经济短期预测模型

陈厚积<sup>1</sup>, 胡晓兵<sup>2</sup>, 邓 希<sup>2</sup>

(1. 明尼苏达大学双城分校文理学院, 明尼阿波利斯 55455; 2. 四川大学机械工程学院, 成都 610065)

**摘要:** 传统的宏观经济短期预测研究通常基于定量或定性预测模型, 而 GMDH 算法兼具定性定量的特点, 建立了基于 GMDH 自回归模型的混合预测模型数学模型来进行宏观经济短期预测, 应用相关的统计数据做实证分析, 将所得结果与传统 GMDH 模型、二次自回归模型进行了比较. 结果表明, 本文提出的混合预测模型具有良好的预测精度, 是一种有效的宏观经济短期预测手段.

**关键词:** 宏观经济短期预测; GMDH; 二次自回归模型

**中图分类号:** C931 **文献标识码:** A **文章编号:** 0490-6756(2020)05-0915-05

## A short-term macroeconomic Forecasting model based on GMDH

CHEN Hou-Ji<sup>1</sup>, HU Xiao-Bing<sup>2</sup>, DENG Xi<sup>2</sup>

(1. School of Arts and Sciences, University of Minnesota Twin Cities, Minneapolis 55455;

2. College of Mechanical Engineering, Sichuan University, Chengdu 610065, China)

**Abstract:** Traditional short-term macroeconomic forecasting research is usually based on quantitative or qualitative forecasting models, while GMDH algorithm has both qualitative and quantitative characteristics. This work proposes a mixed mathematical prediction model to forecast short-term macroeconomic. The model is a combination of GMDH autoregressive model and quadratic autoregressive model. Relevant statistical data are used for empirical analysis, and the results are compared with the traditional GMDH model and the quadratic autoregressive model. The results show that the hybrid forecasting model proposed in this paper has good forecasting accuracy and is an effective means of economic forecasting.

**Keywords:** Short-term macroeconomic forecasts; GMDH; Quadratic autoregression model

## 1 引言

宏观经济预测是一门研究宏观经济系统的发展过程和变动发展趋势的科学. 近年来, 随着我国的经济体制逐步向社会主义市场经济新体制发展, 对宏观经济预测的应用已经十分普遍<sup>[1]</sup>. 计量经济学、时间序列分析法、投入产出分析方法、SAS

预测法和人工神经网络模型<sup>[2]</sup>预测法等已经广泛应用于国民经济分析预测<sup>[3-4]</sup>.

GMDH 自回归模型是自组织数据挖掘(Self-Organization Data Mining)中的一种模型方法, 是基于计算机科学和神经网络的迅速发展而产生和发展的<sup>[5]</sup>. 在信息爆炸和信息不完全的求解条件下, 自组织方法能够有效避免个人主观因素的干

收稿日期: 2020-07-09

作者简介: 陈厚积(1997-), 男, 研究方向为金融经济学、数据分析. E-mail: chen4793@umn.edu

通讯作者: 邓希. E-mail: 1459145424@qq.com

扰<sup>[6]</sup>,在数据挖掘、知识发现、预测、系统建模和模式识别等领域是十分有效的。中国正处于经济结构转型升级的时期,这个经济系统比其他国家更显复杂性,经济指标数据常显示样本量小且受干扰的特点。自组织数据挖掘方法在小样本、受到噪声干扰的经济系统建模预警时比计量经济学方法<sup>[7]</sup>和人工神经网络<sup>[8]</sup>方法更优越<sup>[9]</sup>。

宏观经济短期预测主要是对 GDP、规模以上工业增加值、财政收入、居民消费价格指数等指数进行短期预测<sup>[10]</sup>。本文将传统 GMDH 模型和二次自回归模型进行组合,建立适合经济运行特点的基于 GMDH 的组合模型,应用相关的统计数据做实证分析,能够较为准确地实现对宏观经济的短期预测。

## 2 GMDH 模型原理

GMDH 自回归模型是自组织数据挖掘(Self-Organization Data Mining)中的一种模型方法,它是一种基于遗传进化原理的演化方法,依据给定的准则,从一系列候选模型的集合中挑选出较优的模型<sup>[11]</sup>。GMDH 算法通过遗传、变异和选择,从一个简单的初始模型生成很多具有不断增长复杂度的候选模型,直至模型在观测样本数据上产生过拟合,即最优复杂度为止<sup>[12]</sup>。

在建立 GMDH 模型<sup>[13]</sup>时,将样本数据划分为训练集 A 和检测集 B 两部分,将待预测的数据作为预测集 C。将训练集 A、检测集 B、预测集 C 的数据点个数分别记为  $N_1、N_2、N_3$ 。设共集为  $W = A \cup B$ 。

计算 GMDH 模型的最大可能时滞为:  $m = \sqrt{N_1 + N_2}$ , GMDH 模型参考函数为

$$Y_t = a_0 + a_1 \times Y_{t-1} + a_2 \times Y_{t-2} + \dots + a_m \times Y_{t-m} \quad (1)$$

模型实现步骤如下。

**步骤 1** 将式(1)中每个加法项中的非系数部分看成一个新的自变量,一共产生  $m$  个新的自变量,然后将这  $m$  个自变量作为第一层的输入,  $m$  个自变量两两组合,计算得产生的局部函数的个数为  $M = C_m^2$ ;

**步骤 2** 在训练集 A 和共集 W 上分别使用最小二乘法拟合这些局部函数的系数,从而得到中间模型;

**步骤 3** 利用这些中间模型,计算其对应的训练集 A、共集 W 和预测集 C 上的估计值分别为  $Y1_t、Y2_t、Y3_t$ ;

**步骤 4** 以每个中间模型的绝对抗干扰准则值作为目标函数,函数计算公式为

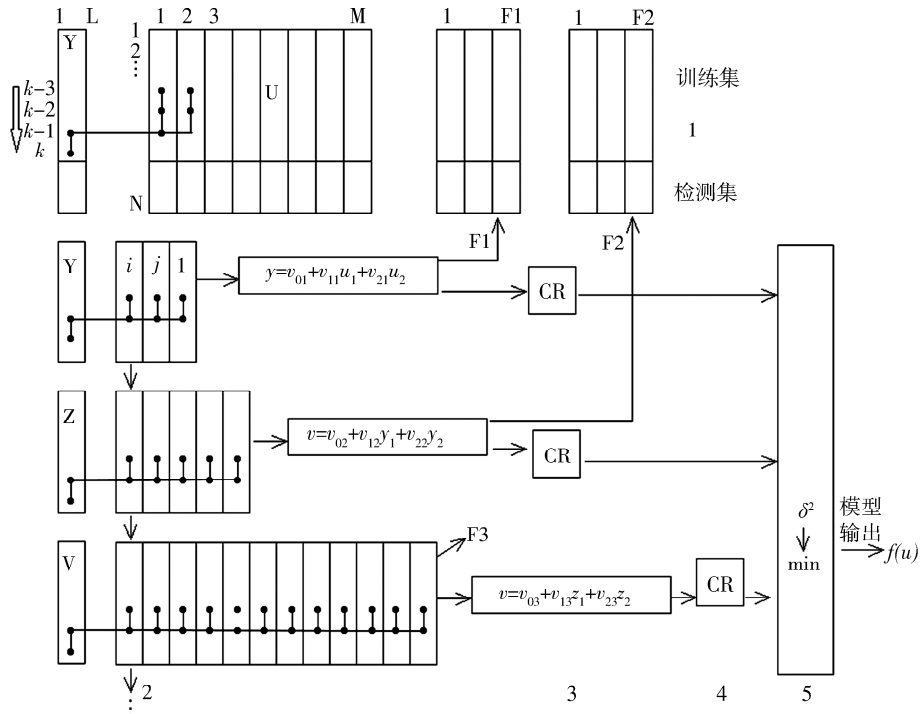


图 1 GMDH 自回归模型流程图  
Fig. 1 Flow chart of GMDH

$$\Delta^2(A) = \sum_{t \in A} (y_t^m(A) - y_t^m(W))^2 \quad (2)$$

**步骤 5** 根据当前层的最小外准则值进行判断, 如果最小外准则值不再减小且当前层数不是第一层, 则停止循环, 跳转到步骤 7, 否则继续步骤 6;

**步骤 6** 淘汰当前层中外准则值较大的  $M - m$  个中间模型, 将剩下  $m$  个中间模型的估计值  $Y_{2t}, Y_{3t}$  保留进入下一层; 重新两两组合, 产生  $M$  个新的局部函数, 跳转到步骤 2;

**步骤 7** 最小外准则值所对应的中间模型就是最优复杂度模型, 估计值  $Y_3$  为需预测的样本值.

GMDH 自回归模型流程图如图 1 所示.

### 3 二次自回归模型原理

使用二次回归模型进行预测的步骤如下:

**步骤 1** 根据预测指标序列中实际数据个数确定滞后项个数  $n$ ,  $n$  的最大值为: 待预测指标序列中实际数据个数  $N$  的平方根  $\sqrt{N}$ .

**步骤 2** 根据滞后项个数  $n$ , 建立不同的模型结构, 曲线公式为

$$y_t = \alpha + \beta_{t-1}(y_{t-1})^{k_1} + \beta_{t-2}(y_{t-2})^{k_2} + \dots + \beta_{t-n}(y_{t-n})^{k_n}, k_n = 1, 2 \quad (3)$$

式中,  $\alpha, \beta$  为待估参数,  $t$  从 1 取到预测期  $z$ , 根据公式计算出一系列的预测值点, 构成曲线. 如果滞后项  $n$  大于  $t$ , 则舍去相应  $t$  值对应的数据项.

**步骤 3** 估计参数  $\alpha$  和  $\beta$  的值, 采用最小二乘估计, 根据式(4).

$$Q = \sum (Y_t - y_t)^2 = \sum e_t^2 = \sum (Y_t - \alpha - \beta_{t-1}y_{t-1} - \beta_{t-2}y_{t-2} - \dots - \beta_{t-n}y_{t-n}) \quad (4)$$

求解目标是使  $Q$  值极小. ( $Y_t$  为实际值) 求  $Q$  对  $\alpha$  和  $\beta$  的偏导, 令其为零, 即

$$\frac{\partial Q}{\partial \alpha} = 0, \frac{\partial Q}{\partial \beta} = 0 \quad (5)$$

根据公式(5) 求出符合要求的待估参数  $\alpha$  和  $\beta$ .

**步骤 4** 检验模型, 计算绝对误差平方和, 即残差, 如式(6),

$$\sum e_t^2 = \sum (Y_t - \alpha - \beta_{t-1}y_{t-1} - \beta_{t-2}y_{t-2} - \dots - \beta_{t-n}y_{t-n}) \quad (6)$$

残差的值越接近于零, 模型越好. 二次自回归模型流程如图 2 所示.

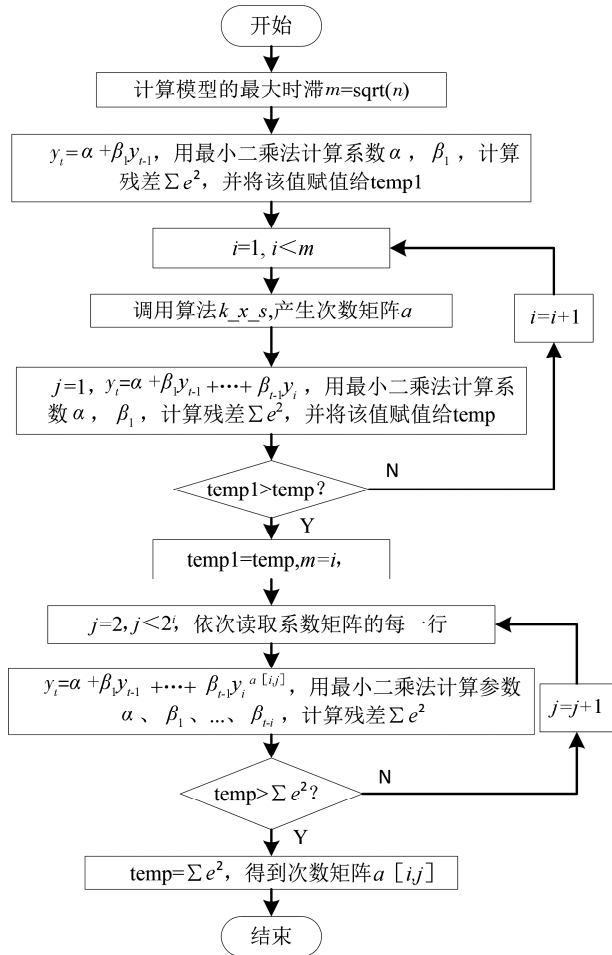


图 2 二次自回归模型流程图  
Fig. 2 Flow chart of quadratic autoregressive model

### 4 组合 GMDH 模型实证分析

应用 GMDH 模型与二次自回归模型模型进行宏观经济短期预测, 以市 GDP 值为例, 表 1 是某市 2006 年 1 月到 2008 年 4 月的市 GDP 值.

表 1 2006 年 1 月~2008 年 4 月某市 GDP 值  
Tab. 1 The city GDP from 2006. 1 to 2008. 4

序号	时间	指标实际值/千万元
1	2006 年 1 月	2 872. 80
2	2006 年 2 月	2 860. 30
3	2006 年 3 月	2 896. 60
4	2006 年 4 月	2 873. 70
5	2006 年 5 月	2 942. 90
6	2006 年 6 月	2 947. 40
7	2006 年 7 月	2 966. 00
8	2006 年 8 月	2 980. 80
9	2006 年 9 月	3 037. 30

(续表 1)

序号	时间	指标实际值/千万元
10	2006年10月	3 089.70
11	2006年11月	3 125.80
12	2006年12月	3 175.50
13	2007年1月	3 253.30
14	2007年2月	3 267.60
15	2007年3月	3 264.30
16	2007年4月	3 289.10
17	2007年5月	3 259.40
18	2007年6月	3 267.60
19	2007年7月	3 239.10
20	2007年8月	3 226.40
21	2007年9月	3 154.00
22	2007年10月	3 190.40
23	2007年11月	3 249.90
24	2007年12月	3 330.00
25	2008年1月	3 320.00
26	2008年2月	3 340.00
27	2008年3月	3 350.00
28	2008年4月	3 380.00

将 2006 年 1 月到 2007 年 12 月的数据作为样本数据,预测 2008 年 1 月到 2008 年 4 月的值,并将它们与实际值进行比较. GMDH 模型预测值如表 2 所示,二次自回归模型预测值如表 3 所示.

表 2 实际值与 GMDH 模型预测值对比

Tab. 2 Actual value vs. estimates of GMDH

序号	时间	预测值 /千万元	实际值 /千万元	绝对误差 /千万元	相对误差 /%
1	2008年1月	3 372.53	3 320.00	52.53	1.55
2	2008年2月	3 371.52	3 340.00	31.52	0.93
3	2008年3月	3 374.74	3 350.00	24.74	0.73
4	2008年4月	3 375.67	3 380.00	-4.33	0.12

表 3 实际值与二次自回归模型预测值对比

Tab. 3 Actual value vs. estimates of quadratic autoregressive model

序号	时间	预测值 /千万元	实际值 /千万元	绝对误差 /千万元	相对误差 /%
1	2008年1月	3 356.2	3 320.00	36.2	1.1
2	2008年2月	3 362.63	3 340.00	22.63	0.7
3	2008年3月	3 345.09	3 350.00	-4.91	0.1
4	2008年4月	3 332.5	3 380.00	-47.5	1.4

由表 2 和表 3 可知,GMDH 模型预测结果与实际值绝对误差与相对误差控制在 1.6%以内,计算得均方差为 1 095.9,二次自回归模型预测结果与实际值绝对误差与相对误差控制在 1.4%左右,计算得均方差为 1 025.7,具有较好的预测效果,但不同预测结果之间相对误差偏大,将 GMDH 模型预测值 $y_1$ 和二次自回归模型预测值 $y_2$ 以  $y = (y_1 + y_2) / 2$  计算后得到新的组合预测值如表 4 所示.

表 4 实际值与组合模型预测值对比

Tab. 4 Actual value vs. estimates of Hybrid model

序号	时间	预测值 /千万元	实际值 /千万元	绝对误差 /千万元	相对误差 /%
1	2008年1月	3 364.37	3 320.00	44.37	1.32
2	2008年2月	3 367.08	3 340.00	27.08	0.80
3	2008年3月	3 359.92	3 350.00	9.92	0.29
4	2008年4月	3 354.09	3 380.00	-25.91	0.77

由表 4 可知 GMDH 模型与二次自回归模型的组合模型相对误差控制在 1.35%以内,计算均方差为 867.9,有效的减少了均方差,不同模型预测和实际值数据对比如图 3 所示,可以看出组合预测模型曲线整体拟合度更高,预测效果较好.证明本文研究的基于 GMDH 的混合模型预测是有效的且精度较高.

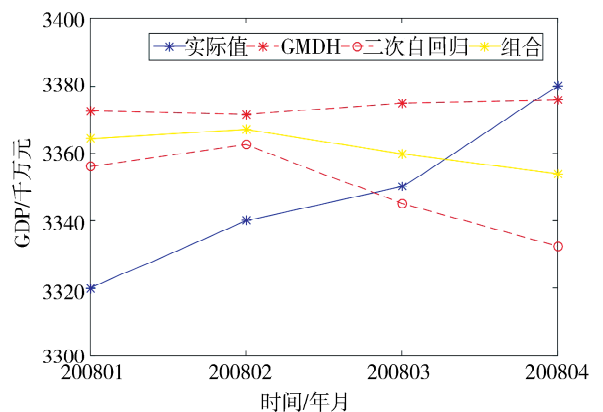


图 3 不同模型预测和实际值数据对比图

Fig. 3 Actual value vs. estimates of different model

## 5 结 论

实验证明 GMDH 模型和二次自回归模型能有效预测短期经济的发展趋势,本文将 GMDH 模型和二次自回归模型进行组合,得到的基于 GMDH 的组合模型能够得到精度较高的预测值,能实

现科学地进行宏观经济短期预测,辅助宏观经济决策部门及时、准确地把握区域经济运行态势。

### 参考文献:

- [1] 张新云. 宏观经济预测理论及其在中国宏观经济中的应用[D]. 济南: 山东大学, 2015.
- [2] 陈鸿. 神经网络在宏观经济预测中的应用研究[D]. 重庆: 重庆大学, 2005.
- [3] 何跃, 马海霞. 基于GMDH的宏观经济景气预测模型及应用[J]. 统计与决策, 2007(6): 52.
- [4] Dendramis Y, Kapetanios G, Marcellino M. A similarity-based approach for macroeconomic forecasting [J]. *J R Stat Soc A Stat*, 2020, 183: 801.
- [5] Mueller J A, Lemke F. Self-organising data mining [M]. Hamburg: Libri, 2000.
- [6] 于姗姗, 董敏. 基于GMDH的成都市宏观经济主要指标预测模型构建[J]. 商场现代化, 2008(31): 280.
- [7] 李子奈. 关于现代计量经济学的研究方法[J]. 清华大学学报: 哲学社会科学版, 1995(3): 29.
- [8] 王志刚. 神经网络在宏观经济预测中的应用研究[J]. 现代经济信息, 2018(1): 18.
- [9] 何跃, 尹静. 基于GMDH的小样本数据预测模型[J]. 统计与决策, 2011(10): 11.
- [10] 刘昕. 宏观经济数据短期统计预测的应用及展望[J]. 中国统计, 2017(9): 6.
- [11] 吉新娜, 贺广福. GMDH模型在青海经济预测系统中的应用[J]. 计算机应用与软件, 2014, 31: 48.
- [12] Zhang M Z, He C Z. D-GMDH: A novel inductive modelling approach in the forecasting of the industrial economy [J]. *Econ Model*, 2013, 30: 514.
- [13] 曹玉洁, 何跃, 贺昌政. 基于R/S分析的GMDH自组织方法在用电量预测中的应用[J]. 软科学, 2009, 23: 17.

### 引用本文格式:

中文: 陈厚积, 胡晓兵, 邓希. 基于GMDH的宏观经济短期预测模型[J]. 四川大学学报: 自然科学版, 2020, 57: 915.

英文: Chen H J, Hu X B, Deng X. A short-term macroeconomic Forecasting model based on GMDH [J]. *J Sichuan Univ; Nat Sci Ed*, 2020, 57: 915.