

摘要

目前的系统辨识方法一般需要在系统输入端加入激励信号,而且多输入多输出系统的在线辨识仍很困难。本文提出一种基于牛顿迭代法的多输入、多输出对象模型迭代辨识方法,模型参数更新的依据是使模型预测输出与全部采样时刻的对象实际输出之间的均方差递减,直到收敛。这种基于全局数据迭代的辨识方法与一般的辨识方法相比,可进行闭环辨识,无需外加激励信号,适用于多输入多输出对象的在线辨识。

(本文首先介绍了在系统辨识领域和电厂锅炉燃烧优化控制领域国内外的研究现状。由于本文提出的模型迭代系统辨识方法是以最小二乘法为基础加以发展变化而来的,本文在其后系统阐述了在系统辨识领域使用最广泛的最小二乘类的系统参数辨识方法。包括最小二乘法的基本原理,最小二乘参数估计的递推算法,还有包括偏差补偿最小二乘法,增广最小二乘法,广义最小二乘法等多种分支最小二乘法。

然后,本文重点介绍了上面提出的基于牛顿迭代法的多输入多输出对象模型迭代辨识的新方法。将其应用在一个两输入两输出的已知模型上面,得到令人满意的结果。基于采用辐射能反馈信号的火电机组负荷控制系统的思想,将其分别应用在一个 200MW 火电机组的现场数据上,对其辨识结果进行了分析,同时还有在单输入单输出模型的给粉-辐射能过程模型辨识结果,结果同样是令人满意的。由于在对对象模型进行系统辨识时,要进行一些假设,如模型的结构和阶次,本文对不同结构和阶次的模型进行了辨识试验,得到了一些有价值的结论。

最后,对本文工作进行总结,并对进一步的研究工作提出了建议和构想。

关键词: 系统辨识; 多输入多输出系统; 最小二乘法;

Abstract

Almost all of the normal identification methods need stimulating signal added into the input of the system. And their behavior in online identification of multi-input multi-output system is not good. The paper developed an iteration identification method for multi-input multi-output system based on the Newton iteration method. This identification method based on global data iteration. It updates the model parameters by minimizing the mean square deviation between the predicted output by the model and the real output of the object during the whole sampling time until the iteration converges. This method can deal with the closed loop identification and be free from stimulating signal.

As first, this paper introduced the main research fruits and results of many scholars all over the world in the fields of the system identification and the optimizing control of fossil-fired power generation unit. For the iteration identification method in this paper based on the least square method, the paper introduced the theory of the least square method sequentially. This part included the principle of the least square method, the recursive algorithm of the least square estimation of parameter, the deviation compensation least square method, the augmented least square method, the generalized least square method, and so on..

In the third section, the paper introduced the new iteration identification method for multi-input multi-output system based on the Newton iteration method. The calculating results for a two-input two-output model, a single-input single-output system, a 300MW and a 200MW fossil-fired power generation units showed good performance of the strategy. The paper also practiced this method on variable constructions and orders and obtained some valuable conclusions.

华中科技大学硕士学位论文

Finally, summary of the overall work and some advice on further research are given.

Keywords: system identification, multi-input multi-output system, the least square method

1 绪论

1.1 课题背景

当前世界三大问题的其中两个是资源和环境。特别是化石燃料资源，例如地底下的煤，海底的油气，都是地久天长积累形成的，它们对国民经济的民展起着极其重要的作用。

在已知的各种能源资源中，煤炭资源最为丰富。根据第 15 届世界能源会议提供的资料，独联体、美国和中国是世界三大煤炭国家，他们占世界总量 57.4%。按 1994 年全世界煤炭消耗量 44.67 亿吨计算，可开采的年限为 233 年，而中国仅为 92 年。事实上，中国的煤贮藏量被低估了，1995 年探明中国煤的保有量为 10087.08 亿吨。在电力生产中，化石燃烧发电是主要的，特别是燃煤电站在当今世界上仍占主导地位，约占 38%，由此可见，煤炭的生产和消费和火力发电的关系是非常密切的。众所周知，我国的火力发电厂绝大多数是燃煤电厂。^[1]

我国的煤炭产量，1987 年为 9.28 亿吨。从这一年开始直到现在，总是位居世界首位，1996 年煤炭产量已达 12.24 亿吨。虽然我国的煤炭产量很高，但人口众多，人均拥有量和人均消费量并不高。同时，在能源利用率方面，与工业发达国家比较，我国也是比较落后的，仅 28%~32%，而欧洲共同体为 40%，美国为 51%，日本高达 75%。在锅炉效率和发电效率方面我们也存在较大的差距。例如，我国现有 40 多万台燃煤工业锅炉，其平均热效率不超过 60%，它消耗每年煤产量的 1/3；13 万多台燃煤工业窑炉的平均热效率仅 20%~30%，每年消耗 15% 的煤产量，民用炉灶的热效率更低。电站锅炉的热效率可达 90%，尽管如此，电力 1 千瓦小时的标准煤耗约为 385 克，比国外先进水平高 80~90 克，因此，全国燃煤电站发电效率仅 29%，比发达国家低 3.5%~13%。

以上情况表明，我国虽然是产煤大国，但人均拥有量并不高，而且能

华中科技大学硕士学位论文

源利用率很低，浪费较大，国民经济效益不大。其原因是多方面的，改进的方法也是多种多样的，但是从煤的利用角度来看，提高煤的燃烧效率是相当重要的。就现有的情况来看，如果从对锅炉运行的优化控制方面来着手进行改进，是一种见效快效果好的改进方法。

辨识、状态估计和控制理论是现代控制论三个互相渗透的领域。辨识和状态估计离不开控制理论的支持，控制理论的应用又几乎不能没有辨识和状态估计技术。随着控制过程复杂性的提高，控制理论的应用日益广泛。但是，它的实际应用不能脱离被控对象的数学模型。然而在多数情况下，被控对象的数学模型是不知道的，或在正常运行期间模型的参数可能发生变化，因此利用控制理论去解决实际问题时，首先需要建立被控对象的数学模型。比如，为了分析、设计一个控制系统，需要建立控制对象（生产设备或生产过程）的数学模型；对生物规律、药物反应或社会经济等问题进行定量分析时，也需要建立相应的数学模型。总之，充分掌握研究对象的运动规律，在表征它们的因果关系时，确立对应的数学模型，这是控制理论能否成功地用于实际的关键之一。辨识正是适应这一需要而形成的一门学科，它的理论正在日趋成熟，它的实际应用已遍及许多领域。

我们谈到辨识，就不能不谈到系统的概念。就工业过程或装置而言，所谓系统就是由元件、部件朴素联系而组成的具有某种特定功能的整体。这种系统，小的如一个晶体管、一个运算放大器、一个电机；大一些的如一个伺服机构、一个调节器、一个惯性导航平台；再大一些如一条生产线、一个电力调节网等。关于系统的概念，已经扩展到其它各个学科领域，是一个应用很广泛的术语。存在有工程系统，生物系统，经济系统，还有社会系统等等。对一般的系统而言，它可以定义为由相互联系、相互制约、相互作用的各个部分组成的，具有一定整体功能和综合行为的统一体^{[2]、[3]}。

不同的系统，对应着不同的学科领域。每个学科领域都要研究这一领域所包含的所有系统内部的以及与它相联系的外部因素之间比较精确的定量关系，即数学模型。所谓系统的数学模型，就是利用数学结构（表格、

图形以及各种数学方程：代数方程、积分方程、微分方程、它们的方程组、传递函数、状态方程、差分方程等）来反映系统内部之间、内部与某些外部因素之间的精确的定量关系。建立数学模型有两种方法：一种是从基本物理定律，即利用各个专门学科领域提出来的物质和能量的守恒性和连续性原理，以及系统（设备）的结构数据推导出模型。这种方法得出的数学模型称之为机理模型或解析模型；这种建立模型的方法称之为解析法。另一种方法是从系统的运行和试验数据建立系统的模型（模型结构和参数），这种方法称之为系统辨识。

关于系统辨识的定义及辨识的译法，目前还不统一（有的译成识别、测辨等）。Zadeh^[4]曾经给系统辨识一个定义：“系统辨识是在输入输出的基础上，从一类系统中确定一个与所测系统等价的系统。”一般概念，所谓辨识就是通过测取研究对象在人为输入作用下的输出响应，或正常运行时的输入输出数据记录，加以必要的数据处理和数学计算，估计出对象的数学模型。这是因为对象的动态特性被认为必然表现在它的变化的输入输出数据中，辨识只是利用数学方法从数据序列提炼出对象的数学模型而已。

从长远看，解析法仍将是建立模型的重要手段，但这种方法只能用于建立比较简单的系统（白箱）模型；对于大多数系统，过程是很复杂的（黑箱和灰箱），以致用解析模型很难准确描述。经过客观实践检验的基本物理定律，最初都是经过反复实验和观测而建立的。从这一意义上讲，用系统辨识方法建立系统的数学模型，则是分析法建模的基础。由于我们所要研究的系统，越来越复杂，而且千变万化，因此，用系统辨识法建立系统的数学模型，将永远处于重要的基础的地位。因此在实际应用中，将两种建模方法结合起来尽量利用我们对物理过程的认识，将系统的模型结构分成已知的和未知的两部分，然后利用实测数据，将未知部分辨识出来。

如果说系统辨识是认识客观世界，那么对系统按人们的意愿进行控制，就是改造客观世界。由于我们所要控制的系统是比较复杂的，它遭受各种各样的干扰，是一个随机过程，同时它们大部分是一个时变系统，因

此，必须对系统进行在线辨识，同时提出实时的控制律，才能进行有效的控制。从内容上讲，在线辨识和最优控制的结合，就是自适应控制。自适应控制离不开系统辨识。

1.2 国内外研究现状

1.2.1 系统辨识的研究发展与现状

在现代控制理论问世之前，大多数被研究的对象是用传递函数描述的线性单输入单输出定常系统。经典控制理论获得极大的成功并得到了广泛应用的原因之一就在于，能运用十分有效的试验方法确定系统的传递函数。这种试验方法，简单易行，如只要在输入端加入正弦变化信号，并在稳态下测量输入输出之间的振幅与相位的关系即可。这种获得传递函数的方法可以直接用于设计控制系统。从这一意义上讲，系统辨识在经典控制理论中已经广泛使用了。

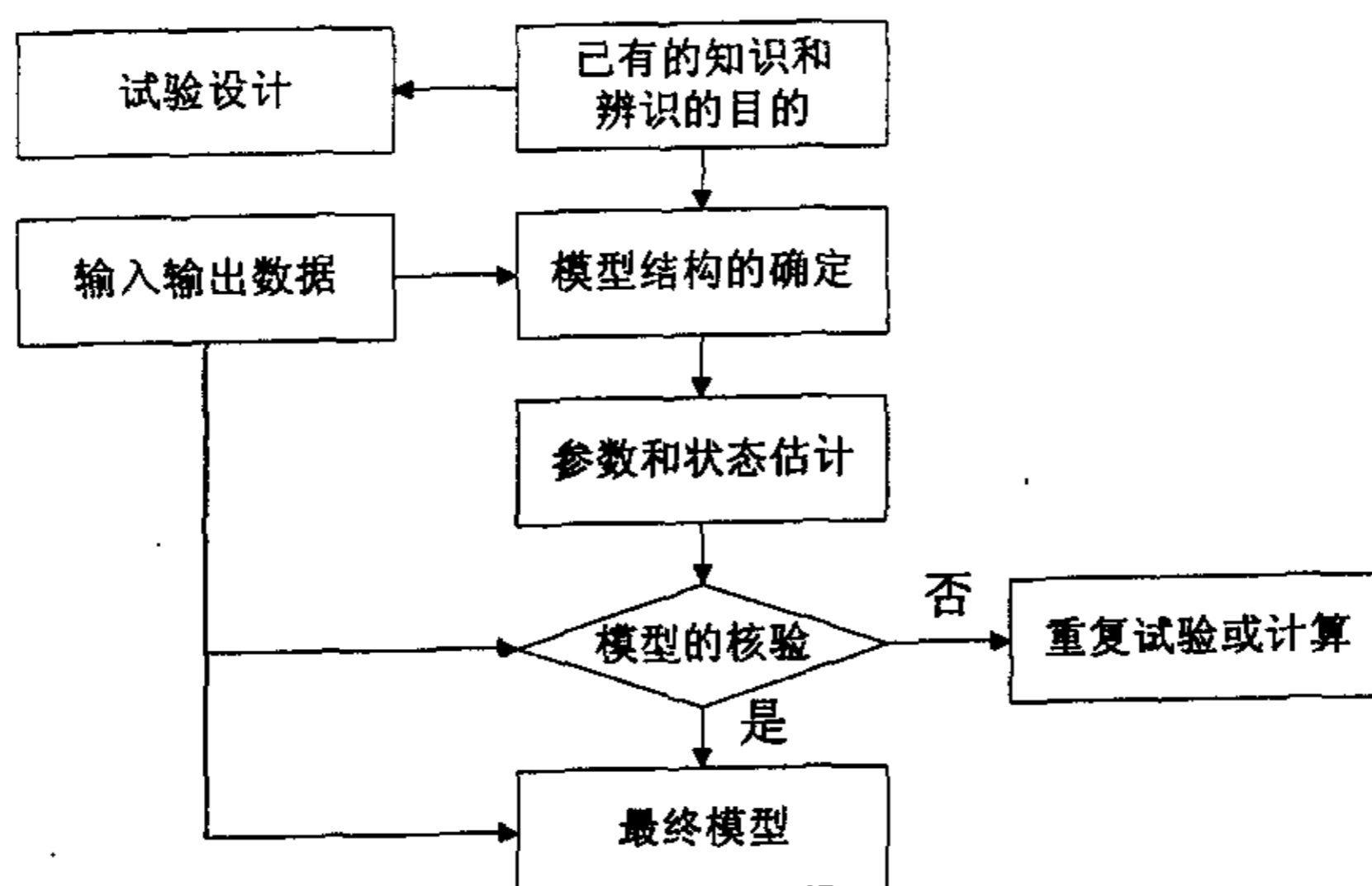


图 1-1 系统辨识的基本内容和步骤

辨识的最终目的是为实际过程设计控制方案。所以辨识系统的数学模型时，首先要根据辨识的目的以及我们对过程已有的认识来确定模型结构，并设计出为了辨识应做的一些试验，在试验进行过程中记录输入输出数据，对模型中的未知部分进行估计。我们所得的数据总是在某种工况

下的数据，在变化的工况下，所估计出来的模型是否仍然适用，就需要用不同工况下的数据，对模型进行核验，不合要求还要重新计算或重新进行试验。因此，辨识的过程是一个迭代过程。这样的做法，最后可能得到一个满意的模型。对于系统辨识的这个基本内容及步骤，可表示为图 1-1。

下面就框图的各部分再进一步做些简单介绍。

1、试验设计

试验设计是系统辨识首先遇到的问题。试验设计的好坏，直接影响系统模型辨识的精度、速度和经济性等。

已有的验前知识、前几次试验结果和辨识的目的对试验设计起重要作用。在安排试验时，首先根据试验对象，确定所要观测的变量。在这些变量中，输入变量可由人们给定，而输出变量要都检测出来。

在许多情况下，为了得到试验设计所必要的知识，必须进行一些所谓预备性试验，它们也是系统辨识的组成部分。可以用一些简单的办法，如阶跃响应法、频率响应法等。通过这些试验，获得系统的信息如下：

- (1) 主要的时间常数；
- (2) 允许的输入信号幅度；
- (3) 过程的非线性与时变性；
- (4) 噪声水平；
- (5) 变量之间的延迟等。

如果系统允许输入扰动信号，如何选择输入信号将是一个很重要的问题。根据不同输入信号的特点，可以应用不同的辨识方法。现在实际中常用阶跃信号、方波、窄脉冲、二进制伪随机噪声等，都是有效的输入信号。

试验的另一个重要问题是采样速度。一般采样速度是根据过程的动态和输入信号特性而定，但是预先精确地确定一个最优采样速度是很复杂的。所以在一般工业装置上收集数据时，可采用实际可能的较短的时间间隔。在数据分析时，可根据需要隔几个数据取一个数据。

试验的时间问题。在正常工况下，辨识精度取决于试验的时间长短，愈长愈好。在试验中，如果过程变化复杂，试验时间不够长可能会得出完

全错误的模型。

2、模型结构的分类及确定

模型结构形式的选择，主要取决于模型的应用目的及精度要求。要在模型使用精度和模型的复杂性之间进行权衡。

常用的模型有以下几类：非参数的-参数的；连续的-离散的；时域的-频域的；单变量的-多变量的；静态的-动态的；线性的-非线性的；时变的-非时变的；确定性的-随机的；集中参数的-分布参数的等等。

数学模型用数据表格或图形表示，即为非参数模型，例如系统的阶跃响应、脉冲响应和频率响应的记录图形（如 Bode 图）等。如果系统辨识的目的是分析系统的特性，系统的时间常数、截止频率等，可采用非参数模型。如果系统辨识的目的是为了对系统进行精确控制，那么要采用参数模型。

一般说来，从参数模型转变成非参数模型（图表和特征值）是不困难的，因为这是已知方程的求解和表征其性质的问题。但是，其逆过程，即由非参数模型求参数模型，则往往是困难的。因为这是前者的反问题，即根据已知解求原方程的问题。若模型结构已知，则较容易从非参数模型求出参数模型。

根据系统的空间、时间变量的离散化情况，模型又可以分为三类：

(1) 集中参数模型。空间变量是离散的，时间变量是连续的，如常微分方程、代数方程；

(2) 离散时间模型。空间、时间变量都是离散的，如差分方程、代数方程；

(3) 分布参数模型。空间、时间变量都是连续的，如偏微分方程等。

其中分布参数模型一般可以在空间上离散化，可以变成离散时间模型和集中参数模型。

为了获得一个可接受的模型，必须重新考虑模型结构的选择。当然，模型结构的确定，要尽量利用对过程本身的物理意义的认识。若能从物理上得到某种模型结构的话，是很有利的。

对一个模型也常常需要衡量它的近似程度，因为有时一个和数据符合得很好的模型，其结构也可能是错误的，例如阶数不对或存在滞后相差，这时需要通过一些统计方法进行检验。

3、参数和状态的估计方法

模型结构确定后，其中的未知部分就需要通过实验数据去估计。模型的未知部分绝大多数情况是以未知参数出现的，因此，参数估计的方法将是后面要讨论的主要内容之一。

参数估计的要求就是要辨识出来的模型与实际的过程在某种意义上是最“接近”的，“接近”是用一个准则来衡量的，这个准则是衡量在同样的输入时，模型的输出与过程的实际输出之间的大小。例如通常用的参数估计的最小二乘法所依据的准则就是模型输出与过程输出之间的误差平方和最小。

当系统的模型是状态空间模型时，有时还需要对它的状态进行估计，尤其是递推估计。系统的状态估计，也是系统辨识的组成部分。

4、模型的校验

一个模型被辨识出来之后，是否是可靠的，即是否真实地反映出过程的特性，是需要核验的。如何检验一个已得到的模型，至今还没有很好的解决。模型的好坏主要应在控制系统的实际运行中来鉴定。

为了得到可靠的模型，需要进行多次实验。更重要的是用一套数据辨识出来的模型要用另一套数据来验证和修改。如此交错地核对，可以使模型的质量得到很好的验证。

所有拟定的假设，在可能的时候都要进行校验。例如时间不变性、线性性、模型的阶次、正态性以及噪声是否有色等等。另一方面，由于实验中的种种限制，实际上不可能对各项都进行校验，只能根据实验所得资料和对过程的了解，来判断模型是否处于合理的偏差范围之内。

一般所用到的系统辨识方法，包括有经典的辨识方法和现代的辨识方法，其中经典的辨识方法包括有阶跃响应法、脉冲响应法、频率响应法、相关分析法和谱分析法等；现代的辨识方法主要有最小二乘类参数辨识方

法、梯度校正参数辨识方法和极大似然法。另外还有预报误差方法、Bayes方法和模型参考自适应辨识方法等等^[7]。

在实际应用中，又有了很多新的系统辨识方法。如胡勇等提出的ARMA（自回归滑动平均）多参数模型辨识方法^[8]、唐巍等提出的混沌优化辨识方法^[9]、王作宏等提出的基于多层前向神经网络的系统辨识方法^[10]、夏超英提出的基于复数模型的描述的系统辨识方法^[11]、林景栋等提出的拓朴辨识方法^[12]、徐文等的模糊辨识方法^[13]、田松礼等将遗传算法用于电力系统频率特性参数辨识^[14]等等。总的来讲，这些新的辨识方法，都是在如前所述的理论基础上，利用不同的算法，规定不同的迭代收敛规则，来进行辨识的。不同的辨识方法，可能应用在不同的系统中会有不同的效果，所以我们在实际中要尽可能选择对本系统最优的辨识方法。

1.2.2 锅炉控制研究的发展

锅炉控制作为实现锅炉安全经济运行目标的有效手段，担负着锅炉主、辅机参数控制、回路调节、联锁保护、顺序控制、参数报警、异常显示、性能计算、趋势记录和报表输出的功能，已从辅助运行人员监控锅炉运行发展到实现不同程度的设备起停功能、过程控制和联锁保护的综合体系。

由于锅炉对象的复杂性，使得锅炉控制理论非常丰富。20世纪60~70年代，对锅炉控制的研究主要集中在锅炉的动态特性和数学模型，从线性到非线性，从单变量到多变量，从时不变到时变等，都进行了广泛而深入的研究。最具代表性的有：H. Nicholson（1964，英国）对一台燃油锅炉的动态特性进行了试验研究，获得了以状态变量形式表达的模型传递函数；K. J. Astrom等（1972，瑞典）从机理分析和试验研究两个方面，对160MW的锅炉——汽轮机机组模型进行了研究，获得了简化的全面反映机组动态特性的非线性模型；A. Tyssso等（1976，挪威）采用机理分析法对一台船用锅炉的物理过程进行了研究，获得了该锅炉的高阶线性多变量模型，并

在此基础上设计、实施了多变量控制系统，运行结果表明，与传统的反馈控制相比，在系统稳定性和安全性方面均有较大提高。

20 世纪 80 年代，传统的控制策略受到现代控制理论的严峻挑战。由于传统的控制策略隐含着两个前提，一是要求对象的模型精确、不变化，且为线性；二是要求对象的运行环境确定、不变化。然而，实际情况下的控制对象只能近似满足这些条件。随着应用领域的扩大，控制精度和性能要求的提高，必须考虑控制对象的非线性，运行环境的变化等多种不确定因素的影响，才能得到满意的控制效果。在这种实际需求的激励和计算机的高速、小型、大容量、低成本条件下，现代控制策略应运而生，并迅速在锅炉控制的实际应用中得到发展，有代表性的有：最优化控制，自适应控制，预测控制等。P. H. Jenkins 等（1981，英国）根据对锅炉燃烧产物的检测分析，提出了一个能反映燃烧状况的性能指标，通过实时优化锅炉燃烧控制中的两个主要因素：燃料量和送风量，即风燃配比，来达到性能指标的最小化。这种自寻优燃烧控制系统已在一台 100MW 燃油锅炉上得到了成功应用，并有望在 500MW 燃油锅炉上得以实施；H. Nakamura 等（1989，日本）对一台 500MW 超临界火电机组进行了优化控制，仿真及现场运行结果均表明，在机组负荷快速、大幅度变化的情况下，优化控制系统表现出良好的稳定性，而对于机组运行中的各种变化，优化控制系统表现出良好的鲁棒性；I. Kocaarslan（1993，德国）在一台 750MW 燃煤锅炉上进行了自适应控制研究，现场运行结果表明，自适应控制在提高锅炉热效率和改善蒸汽品质方面，效果良好。国内，Z. Y. Yang 等（1990）针对锅炉蒸汽温度过程控制中迟延及参数时变的特性，提出了一种模型参考自适应控制，对 Smith 预估控制器参数进行实时修正，该控制方案在一台 300MW 调峰机组上获得了良好的控制效果；杨平等（1993）采用预测控制方法对电站锅炉的蒸汽压力系统进行了仿真研究。

随着世界各国工业向着大型、连续、综合化发展，控制系统变得越来越复杂，而系统的不确定性（包括对象的不确定性和环境的不确定性）成为控制系统研究中最困难的问题。与此同时，人工智能的研究得到飞速发

华中科技大学硕士学位论文

展，并迅速以渗透到各领域。自动控制与人工智能的结合产生了智能控制。20世纪90年代，国内外不少学者将智能控制用于锅炉系统的建模、仿真、诊断及控制。具体有：K. F. Reinschmidt (1991, 美国) 采用神经网络方法对电站锅炉进行了仿真研究，同时指出神经网络方法在电站锅炉控制中具有广阔的应用前景；G. Irwin 等 (1995, 英国) 采用前馈神经网络方法建立了 200MW 燃油锅炉的非线性模型，并对之进行了离线控制，取得了良好的控制效果；M. Hunt 等 (1994, 美国) 采用人工智能方法对一台 650MW 机组的热效率、污染物排放及维修状况进行了在线诊断，结果使锅炉的热效率得到提高， NO_x 的排放量得到降低，同时还降低了机组的维修费用；R. C. Booth 等 (1998, 美国) 采用神经网络方法对一台 400MW 的锅炉进行了燃烧优化控制，结果锅炉热效率提高 0.5%， NO_x 的排放量降低 20~25%。国内，吕剑虹等 (1995) 采用具有学习能力的智能 PID 控制器进行电站锅炉过热汽温控制，获得了良好的控制品质；赵利敏等 (1998) 采用 BP 神经网络对电站燃煤锅炉进行了燃烧预测。

系统装置方面，锅炉控制系统由初期的模拟仪表、DDZ 型发展到组装式仪表、单回路、多回路数字控制仪表。20世纪70年代后期，随着美国 Honeywell 公司推出第一台分散控制系统，计算机硬件技术、软件组态技术和通讯技术形成了强大的技术优势，使得前期电站中相对独立的控制系统，在数字技术的支持下形成了控制功能分散、监控参数集中，各子系统信号形成紧密联系的整体，可编程控制仪表、多微机控制系统也越来越显示出强大的生命力。

随着计算机技术、CRT 显示技术、通讯技术和自动控制理论应用的迅速发展，为使锅炉控制能统一到机、炉、电控制的高效智能一体化、信息管理与控制集成化中，以及现场总线控制系统 (FCS) 和智能仪表融入到分散控制系统 (DCS) 的新型控制和保护策略的网络自动化中，国内外许多公司在这方面都做了有益的探索。国外，如美国 Foxboro 公司的整体单元机组控制；德国 Siemens 公司的带凝结水节流模块的新机组控制 (NUC)；美国 MCS 公司的全厂协调自动化策略 (DEC-500)；Elsag Bailey 公司的性

能计算和优化模块 (ConVali) 及多变量控制和神经网络控制; 日本日立公司的动态前馈并行控制; 瑞士 ABB 公司的诊断模块 Optimax-MODI 及 SCO 主汽温度优化控制软件包; 美国 Leeds & Northrap 公司的直接能量平衡 (DEB) 协调控制系统等。国内, 则有山西省电力科学研究所和华源程控科技有限公司的电站锅炉燃烧控制的新策略——能量平衡法及锅炉安全能效调控装置。

1.2.3 电站锅炉燃烧及机组负荷控制的现状

电站锅炉燃烧控制的基本任务是调整燃烧率水平, 使之适应外界负荷的需要, 稳定蒸汽压力, 同时保证燃烧过程在安全经济的工况下进行; 机组负荷控制则是接收外部负荷的指令, 根据主要运行参数的偏差, 对锅炉、汽机协调地进行控制, 从而在满足电网负荷要求的同时, 保持主要运行参数的稳定。

电站锅炉燃烧控制系统包括三个相对独立而又紧密联系的子系统: 燃料控制子系统、送风控制子系统和引风控制子系统。通常, 蒸汽压力、烟气氧量和炉膛负压分别被选为燃料控制子系统、送风控制子系统和引风控制子系统的被调量。由于烟气氧量的测量有较大的惯性迟延, 因此燃烧控制的基本方案采用直接保持燃料量与送风量成一定比例关系, 但又考虑到燃烧过程的经济性, 通常采用烟气氧量信号来修正送风量, 亦即采用氧量校正调节。

另外, 有一种先进的燃烧控制方式——双交叉限幅燃烧控制方式在电站锅炉上得以应用, 其主要特点为:

- (1) 稳态时能保证经济空燃比和最佳烟气含氧量;
- (2) 升负荷时, 保证空气量先增加, 燃料量跟随实测空气量;
- (3) 减负荷时, 保证先减燃料量, 空气量跟随实测燃料量;
- (4) 空燃交叉限制解除功能: 当燃料量手动, 送风量自动时, 送风设定值跟随燃料量; 当送风手动, 燃料量自动时, 燃料量设定值跟随送风

量。

双交叉限幅燃烧控制方式先在燃油、燃气的电站锅炉上得以成功应用，后逐步在燃煤电站锅炉上推广应用。

机组负荷控制方式通常分为二类：机炉分别控制方式和机炉协调控制方式。其中，机炉分别控制方式包括锅炉跟随方式和汽机跟随方式两种，且锅炉跟随方式一般用于蓄热能力相对较大的中、小型汽包锅炉机组，汽机跟随方式一般用于单元机组承担基本负荷的场合；机炉协调控制方式包括以锅炉跟随为基础的协调控制方式、以汽机跟随为基础的协调控制方式和综合型协调控制方式。以锅炉跟随为基础的协调控制方式和以汽机跟随为基础的协调控制方式只实现“单向”的协调，而综合型协调控制方式能够实现“双向”的协调，它的典型代表有：直接能量平衡方式（DEB）。

直接能量平衡方式的特点为：汽机控制机组功率；以能量平衡信号作为锅炉侧的负荷指令，以热量信号作其反馈，直接按汽机的能量需要来控制锅炉的输入。其中，能量平衡信号为： $p_s \times \frac{p_1}{p_T}$ ， p_1 为汽机第一级压力，

p_s 为机前压力设定值， p_T 为机前压力。由于能量平衡信号能够正确反映汽机对锅炉能量的需求，且只反映外扰（汽机调门开度变化），而不受锅炉内扰（燃料扰动）的影响，因此它适用于所有运行工况（包括定压和滑压）。

对于电站燃煤锅炉，由于燃料量到蒸汽压力这条通道具有纯延迟、大滞后的特性，致使一些先进控制算法，如 Smith 预估算法，DAHLIN 算法的控制效果不佳，即使采用热量信号（蒸汽流量信号与汽包压力微分信号之和）代表燃料量，与蒸汽压力一道对入炉燃料量进行串级控制，效果也并不理想。

因此，提高电站燃煤锅炉燃烧及机组负荷控制的品质，有效克服锅炉燃料量到蒸汽压力的滞后特性是关键。

基于炉膛辐射能能够迅速、直观地反映炉内燃烧状况，周怀春等（1996）提出采用炉膛断面辐射能控制燃烧的思想^[15]；王满家等（1995）利用炉内火焰辐射能信号的大小对锅炉的燃料、送风进行调节，使锅炉运

行保持热效率最佳状况。

另一个方面，锅炉蒸汽量的大小、热效率的高低均与炉膛温度水平有密切关系。同时，大量研究表明，燃烧产物中大多数的有害物质（包括 CO 和 NO_x ）生成量的大小都跟炉膛温度紧密相关。因此，若引入基于火焰温度图像的炉膛辐射能进行燃烧控制，必将有利于提高锅炉效率、降低污染物的生成。

炉膛火焰检测技术的进步，为基于火焰温度图像的炉膛辐射能检测奠定了基础，也为采用炉膛辐射能信号进行燃烧控制提供了保障。

1.3 本文主要工作

在火力发电厂中，从燃料、供风和给水等进入锅炉到机组输出电力的过程是一个典型的多输入、多输出、强耦合的复杂过程，其优化控制的研究一直是热工自动化工作者几十年来关注的焦点（例如，[16-19]）。以自适应控制为代表的大量新型控制系统在过程控制中成功应用的关键问题是快速、实时在线辨识对象模型。开环测试显然已不适应实时控制与高性能指标控制的要求，尤其是强迫性加入输入信号会造成人为扰动，影响正常的生产过程^[20]。最小二乘法在模型辨识中得到了广泛应用，但目前的研究一般针对单输入、单输出系统（例如，[21]，[22]）。文献[23]针对作为单输入、单输出系统的汽轮机调节系统的模型辨识问题，提出在单机运行时利用负荷调整信号、并列运行时利用同步器给定扰动信号作为系统的激励信号，采用相关辨识方法获取对象的模型参数。由于火电机组时时刻刻均处于不断的调整过程中，文献[23]的意义可能还在于：无需外界激励信号，直接利用处于不断调整中的系统闭环运行条件下的输入、输出信号，可进行有效的模型辨识。本文提出一种基于全局数据迭代的模型辨识方法，应用于双输入、双输出系统，可以很好地解决多输入多输出系统内部存在的耦合问题。不需要加入激励信号，从而避免了输入外部激励信号对系统正常工作的影响。只要数据序列含有辨识所需要的足够信息量，就可

华中科技大学硕士学位论文

以依据预测输出与实际输出之间的偏差，应用牛顿迭代法进行迭代，直到由得到模型的预测输出与实际输出之间十分接近为止。

2 最小二乘系统辨识方法

2.1 最小二乘法的基本原理

2.1.1 最小二乘法的概念

最小二乘法大约是 1795 年高斯在他那著名的星体运动轨道预报研究工作中提出的。后来，最小二乘法就成了估计理论的奠基石。由于最小二乘法原理简单，编制程序也不困难，所以它颇受人们重视，应用相当广泛。

最小二乘法的基本结果有两种形式，一种是经典的一次完成算法；另一种是现代的递推算法。后者更适用于计算机在线辨识，而前者在理论研究方面却更为方便。

设过程的输入输出关系可以描述成如下的最小二乘格式

$$z(k) = \mathbf{h}^T(k)\boldsymbol{\theta} + n(k) \quad (2.1.1)$$

其中， $z(k)$ 是过程的输出量； $\mathbf{h}(k)$ 是可观测的数据向量； $n(k)$ 是均值为零的随机噪声。利用数据序列 $\{z(k)\}$ 和 $\{\mathbf{h}(k)\}$ ，极小化下列准则函数

$$J(\boldsymbol{\theta}) = \sum_{k=1}^L [z(k) - \mathbf{h}^T(k)\boldsymbol{\theta}]^2 \quad (2.1.2)$$

使 $J(\boldsymbol{\theta}) = \min$ 的 $\boldsymbol{\theta}$ 估计值记作 $\hat{\boldsymbol{\theta}}$ ，称作 $\boldsymbol{\theta}$ 参数的最小二乘估计值。

上述基本概念表明，未知模型参数最可能的值是在实际观测值与计算值之累次误差的平方和达到最小值处，所得到的这种模型输出能最好地接近实际过程的输出。

2.1.2 最小二乘问题的提法

设时不变 SISO 动态过程的数学模型为

$$A(z^{-1})z(k) = B(z^{-1})u(k) + n(k) \quad (2.1.3)$$

其中, $u(k)$ 和 $z(k)$ 为过程的输入输出量; $n(k)$ 是噪声; 多项式 $A(z^{-1})$ 和 $B(z^{-1})$ 如 (2.1.4) 所示。现在的问题是如何利用过程的输入输出数据, 确定多项式 $A(z^{-1})$ 和 $B(z^{-1})$ 的系数。

$$\begin{cases} A(z^{-1}) = 1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_{n_a} z^{-n_a} \\ B(z^{-1}) = b_1 z^{-1} + b_2 z^{-2} + \dots + b_{n_b} z^{-n_b} \end{cases} \quad (2.1.4)$$

在叙述怎样解决这类过程的辨识问题之前, 先明确一些基本假设和基本关系是有益的。

首先, 假定模型 (2.1.3) 式的阶次 n_a 和 n_b 已经设定, 且一般有 $n_a > n_b$ 。

当取相同阶次时, 记作 $n = n_a = n_b$ 。

其次, 将模型 (2.1.3) 式写成最小二乘格式

$$z(k) = \mathbf{h}^T(k)\boldsymbol{\theta} + n(k) \quad (2.1.5)$$

式中

$$\begin{cases} \mathbf{h}(k) = [-z(k-1), \dots, -z(k-n_a), u(k-1), \dots, u(k-n_b)]^T \\ \boldsymbol{\theta} = [a_1, a_2, \dots, a_{n_a}, b_1, b_2, \dots, b_{n_b}]^T \end{cases} \quad (2.1.6)$$

对于 $k = 1, 2, \dots, L$, 方程 (2.1.5) 式构成一个线性方程组, 可以把它写成

$$\mathbf{z}_L = \mathbf{H}_L \boldsymbol{\theta} + \mathbf{n}_L \quad (2.1.7)$$

其中

$$\begin{cases} \mathbf{z}_L = [z(1), z(2), \dots, z(L)]^T \\ \mathbf{n}_L = [n(1), n(2), \dots, n(L)]^T \\ \mathbf{H}_L = \begin{bmatrix} \mathbf{h}^T(1) \\ \mathbf{h}^T(2) \\ \dots \\ \mathbf{h}^T(L) \end{bmatrix} = \begin{bmatrix} -z(0) & \dots & -z(1-n_a) & u(0) & \dots & u(1-n_b) \\ -z(1) & \dots & -z(2-n_a) & u(1) & \dots & u(2-n_b) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ -z(L-1) & \dots & -z(L-n_a) & u(L-1) & \dots & u(L-n_b) \end{bmatrix} \end{cases} \quad (2.1.8)$$

另外, 设模型 (2.1.3) 式的噪声 $n(k)$ 完全可用一阶和二阶统计矩描述,

即设它的均值和协方差阵为

$$\left\{ \begin{array}{l} E\{\mathbf{n}_L\} = \begin{bmatrix} E\{n(1)\} \\ E\{n(2)\} \\ \dots \\ E\{n(L)\} \end{bmatrix} = \mathbf{0} \\ Cov\{\mathbf{n}_L\} = E\{\mathbf{n}_L \mathbf{n}_L^T\} = \begin{bmatrix} E\{n^2(1)\} & E\{n(1)n(2)\} & \dots & E\{n(1)n(L)\} \\ E\{n(2)n(1)\} & E\{n^2(2)\} & \dots & E\{n(2)n(L)\} \\ & \dots & \dots & \dots \\ E\{n(L)n(1)\} & E\{n(L)n(2)\} & \dots & E\{n^2(L)\} \end{bmatrix} \cong \Sigma_n \end{array} \right. \quad (2.1.9)$$

必须指出, 在推导最小二乘法的结果时, 并不需要考虑噪声 $n(k)$ 的统计特性。但是, 在评价最小二乘估计的性质时, 则必须进一步假设噪声 $n(k)$ 是不相关的, 而且是同分布的随机变量。简单地说, 必须假设 $\{n(k)\}$ 是白噪声序列, 即

$$\begin{cases} E\{\mathbf{n}_L\} = \mathbf{0} \\ Cov\{\mathbf{n}_L\} = \sigma_n^2 \mathbf{I} \end{cases} \quad (2.1.10)$$

其中, σ_n^2 为噪声 $n(k)$ 的方差。在评价最小二乘估计量的某些统计性质时, 有时还要假设噪声 $n(k)$ 服从正态分布。此外, 还认为噪声 $n(k)$ 和输入 $u(k)$ 是不相关的, 即

$$E\{n(k)u(k-l)\} = 0, \quad \forall k, l \quad (2.1.11)$$

最后, 如何选择记忆长度或称数据长度 L 也是需要考虑的问题。显然, 联立方程组 (2.1.7) 式具有 L 个方程, 包含 $(n_a + n_b)$ 个未知数。如果 $L < (n_a + n_b)$, 方程的个数少于未知数个数, 模型参数 θ 不能唯一确定, 这种情况一般可以不去讨论它。如果 $L = (n_a + n_b)$, 则只有当 $\mathbf{n}_L = \mathbf{0}$ 时, θ

才有唯一的确定解，这也不是现在所要研究的问题。当 $\mathbf{n}_L \neq \mathbf{0}$ 时，只有取 $L > (n_a + n_b)$ ，才有可能确定一个“最优”的模型参数 θ ，而且为了保证辨识的精度， L 必须充分大。

以上这些基本假设对本文以后的讨论原则上也是适用的。

2.2 最小二乘问题的解

考虑模型 (2.1.5) 式的辨识问题，其中 $z(k)$ 和 $\mathbf{h}(k)$ 都是可观测的数据， θ 是待估计参数，准则函数取

$$J(\theta) = \sum_{k=1}^L \Lambda(k) [z(k) - \mathbf{h}^T(k)\theta]^2 \quad (2.2.1)$$

其中， $\Lambda(k)$ 称为加权因子，对所有的 k ， $\Lambda(k)$ 都必须为正数。引进加权因子的目的是为了便于考虑观测数据的可信度。如果有理由认为现时刻的数据比过去时刻的数据可靠，那么现在时刻的加权值就大于过去时刻的加权值。比如，可选 $\Lambda(k) = \mu^{L-k}$ ， $0 < \mu < 1$ 。当 $k=1$ 时， $\Lambda(1) = \mu^{L-1} \ll 1$ ；当 $k=L$ 时， $\Lambda(L) = 1$ ，这就体现了对不同时时刻的数据给予了不同程度的信任。 $\Lambda(k)$ 的选择多少取决于人的主观因素，并无一般规律可循。在实际应用中，如果对象是线性时不变的过程，或者数据的可信度还难以肯定的话，则可以简单地选择 $\Lambda(k) = 1, \forall k$ 。

根据 (2.1.8) 式的定义，准则函数 $J(\theta)$ 可写成二次型的形式

$$J(\theta) = (\mathbf{z}_L - \mathbf{H}_L \theta)^T \Lambda_L (\mathbf{z}_L - \mathbf{H}_L \theta) \quad (2.2.2)$$

其中， Λ_L 为加权阵，一般是正定的对角矩阵，它与加权因子 $\Lambda(k)$ 的关系是

$$\Lambda_L = \begin{bmatrix} \Lambda(1) & & & \mathbf{0} \\ & \Lambda(2) & & \\ & & \dots & \\ \mathbf{0} & & & \Lambda(L) \end{bmatrix} \quad (2.2.3)$$

显然，(2.2.2) 式中的 $\mathbf{H}_L \boldsymbol{\theta}$ 代表模型的输出，或者说是过程的输出预报值。因此 $J(\boldsymbol{\theta})$ 可以被看作用来衡量模型输出与实际过程输出的接近情况。极小化 $J(\boldsymbol{\theta})$ ，求得参数 $\boldsymbol{\theta}$ 的估计值将使模型的输出最好地预报过程的输出。

设 $\hat{\boldsymbol{\theta}}_{WLS}$ 使得 $J(\boldsymbol{\theta})|_{\hat{\boldsymbol{\theta}}_{WLS}} = \min$ ，则有

$$\left. \frac{\partial J(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right|_{\hat{\boldsymbol{\theta}}_{WLS}} = \left. \frac{\partial}{\partial \boldsymbol{\theta}} (\mathbf{z}_L - \mathbf{H}_L \boldsymbol{\theta})^T \Lambda_L (\mathbf{z}_L - \mathbf{H}_L \boldsymbol{\theta}) \right|_{\hat{\boldsymbol{\theta}}_{WLS}} = \mathbf{0}^T \quad (2.2.4)$$

展开之，并运用如下两个向量微分公式

$$\begin{cases} \frac{\partial}{\partial \mathbf{x}} (\mathbf{a}^T \mathbf{x}) = \mathbf{a}^T \\ \frac{\partial}{\partial \mathbf{x}} (\mathbf{x}^T \mathbf{A} \mathbf{x}) = 2\mathbf{x}^T \mathbf{A}, \mathbf{A} \text{ 为对称阵} \end{cases} \quad (2.2.5)$$

得

$$(\mathbf{H}_L^T \Lambda_L \mathbf{H}_L) \hat{\boldsymbol{\theta}}_{WLS} = \mathbf{H}_L^T \Lambda_L \mathbf{z}_L \quad (2.2.6)$$

上式称为正则方程。当 $\mathbf{H}_L^T \Lambda_L \mathbf{H}_L$ 是正则矩阵时，有

$$\hat{\boldsymbol{\theta}}_{WLS} = (\mathbf{H}_L^T \Lambda_L \mathbf{H}_L)^{-1} \mathbf{H}_L^T \Lambda_L \mathbf{z}_L \quad (2.2.7)$$

另外

$$\left. \frac{\partial^2 J(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2} \right|_{\hat{\boldsymbol{\theta}}_{WLS}} = 2(\mathbf{H}_L^T \Lambda_L \mathbf{H}_L) \quad (2.2.8)$$

因 Λ_L 是正定矩阵，故 $\mathbf{H}_L^T \Lambda_L \mathbf{H}_L$ 也是正定矩阵，即

$$\left. \frac{\partial^2 J(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2} \right|_{\hat{\boldsymbol{\theta}}_{WLS}} > 0 \quad (2.2.9)$$

所以满足 (2.2.7) 式 $\hat{\boldsymbol{\theta}}_{WLS}$ 使得 $J(\boldsymbol{\theta})|_{\hat{\boldsymbol{\theta}}_{WLS}} = \min$ ，并且 $\hat{\boldsymbol{\theta}}_{WLS}$ 是唯一的。

通过极小化 (2.2.2) 式计算 $\hat{\theta}_{WLS}$ 的方法称作加权最小二乘法, 对应的 $\hat{\theta}_{WLS}$ 称为加权最小二乘估计值。如果加权阵取 $\Lambda_L = \mathbf{I}$, 则 (2.2.7) 式退化

$$\hat{\theta}_{LS} = (\mathbf{H}_L^T \mathbf{H}_L)^{-1} \mathbf{H}_L^T \mathbf{z}_L \quad (2.2.10)$$

这时的 $\hat{\theta}_{LS}$ 简称最小二乘估计值, 对应的方法叫做最小二乘法。它是加权最小二乘法的一种特例。

当获得一批数据之后, 利用 (2.2.7) 或 (2.2.10) 式可一次求得相应参数估计值, 这样的处理问题的方法就称作一次完成算法。它在理论研究方面有许多方便之处, 但在计算方面要碰到矩阵求逆的困难。当矩阵的维数增加时, 矩阵求逆运算的计算量将急剧增加, 这会给计算机的计算速度和存储量带来负担, 因此有时也可用高斯消元法直接解正则方程 (2.2.6) 式, 以便更快地求得参数估计值。但是, 更为实用的方法还是设法把 (2.2.7) 式化成递推计算的形式, 这样便于在线辨识, 而且大大减少了数据贮存, 节省了计算机的内存。后面将会详细讨论这种递推算法。另外, 一次完成算法要求 $\mathbf{H}_L^T \Lambda_L \mathbf{H}_L$ 必须是正则矩阵 (可逆矩阵), 其充分必要条件是过程的输入信号必须是 $2n$ 阶持续激励信号 (Åström and Bohlin, 1965), 即要求

$$\bar{\mathbf{U}}_L^T \bar{\mathbf{U}}_L > 0 \quad (2.2.11)$$

其中

$$\left\{ \begin{array}{l} \bar{\mathbf{U}}_L = [\mathbf{F}\mathbf{u}_L, \mathbf{F}^2\mathbf{u}_L, \dots, \mathbf{F}^{2n}\mathbf{u}_L] \\ \mathbf{u}_L = [u(1), u(2), \dots, u(L)]^T \\ \mathbf{F} = \begin{bmatrix} 0 & & \mathbf{0} \\ 1 & \dots & \\ & \dots & \dots \\ \mathbf{0} & & 1 & 0 \end{bmatrix} \\ n = \max(n_a, n_b) \end{array} \right. \quad (2.2.12)$$

上述条件称作开环可辨识性条件。它意味着辨识所用的输入信号不能随意

选择，否则可能造成不可辨识。目前常用的信号有：

- (1) 随机序列（如白噪声）；
- (2) 伪随机序列（如 M 序列或逆 M 序列）；
- (3) 离散序列，通常指对含有 n 种频率（各频率不能满足整数倍关系）的正弦组合信号进行采样处理获得的离散序列。

例如，我们考虑如下仿真对象

$$z(k) - 1.5z(k-1) + 0.7z(k-2) = u(k-1) + 0.5u(k-2) + v(k) \quad (2.2.13)$$

其中， $v(k)$ 为服从 $N(0,1)$ 分布的白噪声。输入信号 $u(k)$ 采用 4 阶 M 序列，幅度为 1。选择如下的辨识模型

$$z(k) + a_1z(k-1) + a_2z(k-2) = b_1u(k-1) + b_2u(k-2) + v(k) \quad (2.2.14)$$

按 (2.1.8) 式构造 \mathbf{z}_L 和 \mathbf{H}_L ；数据长度取 $L=400$ ；加权矩阵取 $\Lambda_L = \mathbf{I}$ ；利用 (2.2.10) 式计算参数估计值，如果如表 2-1 所示。

表 2-1 一次完成算法的计算结果

参数	a_1	a_2	b_1	b_2
真值	-1.5	0.7	1.0	0.5
估计值	-1.52	0.720	0.946	0.566

2.3 最小二乘参数估计的递推算法

上节已经给出了最小二乘一次完成算法，但具体使用时不仅占用内存量大，而且不能用于在线辨识。解决这个问题的办法是把它化成递推算法。递推算法可以概括成

$$\text{新的估计值 } \hat{\boldsymbol{\theta}}(k) = \text{老的估计值 } \hat{\boldsymbol{\theta}}(k-1) + \text{修正项} \quad (2.3.1)$$

新的估计值 $\hat{\boldsymbol{\theta}}(k)$ 是在老的估计值 $\hat{\boldsymbol{\theta}}(k-1)$ 的基础上修正而成。这样不仅可以减少计算量和贮存量，而且能实现在线实时辨识。

2.3.1 依观测次序的递推算法

依观测次序的递推算法就是每获得一次新的观测数据就修正一次参数估计值，随着时间的推移，便能获得满意的辨识结果。

首先，将 (2.2.7) 式一次完成算法写成

$$\begin{aligned}\hat{\boldsymbol{\theta}}_{WLS} &= (\mathbf{H}_L^T \boldsymbol{\Lambda}_L \mathbf{H}_L)^{-1} \mathbf{H}_L^T \boldsymbol{\Lambda}_L \mathbf{z}_L \cong \mathbf{P}(L) \mathbf{H}_L^T \boldsymbol{\Lambda}_L \mathbf{z}_L \\ &= \left[\sum_{i=1}^L \Lambda(i) \mathbf{h}(i) \mathbf{h}^T(i) \right]^{-1} \left[\sum_{i=1}^L \Lambda(i) \mathbf{h}(i) z(i) \right]\end{aligned}\quad (2.3.2)$$

定义

$$\begin{cases} \mathbf{P}^{-1}(k) = \sum_{i=1}^k \Lambda(i) \mathbf{h}(i) \mathbf{h}^T(i) \cong \mathbf{H}_k^T \boldsymbol{\Lambda}_k \mathbf{H}_k \\ \mathbf{P}^{-1}(k-1) = \sum_{i=1}^{k-1} \Lambda(i) \mathbf{h}(i) \mathbf{h}^T(i) \cong \mathbf{H}_{k-1}^T \boldsymbol{\Lambda}_{k-1} \mathbf{H}_{k-1} \end{cases}\quad (2.3.3)$$

其中

$$\begin{cases} \mathbf{H}_k = \begin{bmatrix} \mathbf{h}^T(1) \\ \mathbf{h}^T(2) \\ \dots \\ \mathbf{h}^T(k) \end{bmatrix}, \boldsymbol{\Lambda}_k = \begin{bmatrix} \Lambda(1) & & & \mathbf{0} \\ & \Lambda(2) & & \\ & & \dots & \\ \mathbf{0} & & & \Lambda(k) \end{bmatrix} \\ \mathbf{H}_{k-1} = \begin{bmatrix} \mathbf{h}^T(1) \\ \mathbf{h}^T(2) \\ \dots \\ \mathbf{h}^T(k-1) \end{bmatrix}, \boldsymbol{\Lambda}_{k-1} = \begin{bmatrix} \Lambda(1) & & & \mathbf{0} \\ & \Lambda(2) & & \\ & & \dots & \\ \mathbf{0} & & & \Lambda(k-1) \end{bmatrix} \end{cases}\quad (2.3.4)$$

由 (2.3.3) 式可得

$$\begin{aligned}\mathbf{P}^{-1}(k) &= \sum_{i=1}^{k-1} \Lambda(i) \mathbf{h}(i) \mathbf{h}^T(i) + \Lambda(k) \mathbf{h}(k) \mathbf{h}^T(k) \\ &= \mathbf{P}^{-1}(k-1) + \Lambda(k) \mathbf{h}(k) \mathbf{h}^T(k)\end{aligned}\quad (2.3.5)$$

置

$$\mathbf{z}_{k-1} = [z(1), z(2), \dots, z(k-1)]^T \quad (2.3.6)$$

则

$$\begin{aligned}\hat{\boldsymbol{\theta}}(k-1) &= (\mathbf{H}_{k-1}^T \boldsymbol{\Lambda}_{k-1} \mathbf{H}_{k-1})^{-1} \mathbf{H}_{k-1}^T \boldsymbol{\Lambda}_{k-1} \mathbf{z}_{k-1} \\ &= \mathbf{P}(k-1) \left[\sum_{i=1}^{k-1} \Lambda(i) \mathbf{h}(i) z(i) \right]\end{aligned}\quad (2.3.7)$$

于是有

$$\mathbf{P}^{-1}(k-1) \hat{\boldsymbol{\theta}}(k-1) = \sum_{i=1}^{k-1} \Lambda(i) \mathbf{h}(i) z(i) \quad (2.3.8)$$

令

$$\mathbf{z}_k = [z(1), z(2), \dots, z(k)]^T \quad (2.3.9)$$

并利用 (2.3.5) 和 (2.3.8) 式, 可得

$$\begin{aligned}\hat{\boldsymbol{\theta}}(k) &= (\mathbf{H}_k^T \boldsymbol{\Lambda}_k \mathbf{H}_k)^{-1} \mathbf{H}_k^T \boldsymbol{\Lambda}_k \mathbf{z}_k = \mathbf{P}(k) \left[\sum_{i=1}^k \Lambda(i) \mathbf{h}(i) z(i) \right] \\ &= \mathbf{P}(k) [\mathbf{P}^{-1}(k-1) \hat{\boldsymbol{\theta}}(k-1) + \Lambda(k) \mathbf{h}(k) z(k)] \\ &= \mathbf{P}(k) \{ [\mathbf{P}^{-1}(k) - \Lambda(k) \mathbf{h}(k) \mathbf{h}^T(k)] \hat{\boldsymbol{\theta}}(k-1) + \Lambda(k) \mathbf{h}(k) z(k) \} \\ &= \hat{\boldsymbol{\theta}}(k-1) + \mathbf{P}(k) \mathbf{h}(k) \Lambda(k) [z(k) - \mathbf{h}^T(k) \hat{\boldsymbol{\theta}}(k-1)]\end{aligned}\quad (2.3.10)$$

引进增益矩阵 $\mathbf{K}(k)$, 定义为

$$\mathbf{K}(k) = \mathbf{P}(k) \mathbf{h}(k) \Lambda(k) \quad (2.3.11)$$

则式写成

$$\hat{\boldsymbol{\theta}}(k) = \hat{\boldsymbol{\theta}}(k-1) + \mathbf{K}(k) [z(k) - \mathbf{h}^T(k) \hat{\boldsymbol{\theta}}(k-1)] \quad (2.3.12)$$

进一步把 (2.3.5) 式写成

$$\mathbf{P}(k) = [\mathbf{P}^{-1}(k-1) + \Lambda(k) \mathbf{h}(k) \mathbf{h}^T(k)]^{-1} \quad (2.3.13)$$

为了避免矩阵求逆运算, 利用矩阵反演公式

$$(\mathbf{A} + \mathbf{C}\mathbf{C}^T)^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{C} (\mathbf{I} + \mathbf{C}^T \mathbf{A}^{-1} \mathbf{C})^{-1} \mathbf{C}^T \mathbf{A}^{-1} \quad (2.3.14)$$

可将 (2.3.13) 式演变成

$$\begin{aligned} \mathbf{P}(k) &= \mathbf{P}(k-1) - \mathbf{P}(k-1)\mathbf{h}(k)\mathbf{h}^T(k)\mathbf{P}(k-1) \left[\mathbf{h}^T(k)\mathbf{P}(k-1)\mathbf{h}(k) + \frac{1}{\Lambda(k)} \right]^{-1} \\ &= \left[\mathbf{I} - \frac{\mathbf{P}(k-1)\mathbf{h}(k)\mathbf{h}^T(k)}{\mathbf{h}^T(k)\mathbf{P}(k-1)\mathbf{h}(k) + \Lambda^{-1}(k)} \right] \mathbf{P}(k-1) \end{aligned} \quad (2.3.15)$$

将上式代入 (2.3.11) 式, 整理后有

$$\mathbf{K}(k) = \mathbf{P}(k-1)\mathbf{h}(k) \left[\mathbf{h}^T(k)\mathbf{P}(k-1)\mathbf{h}(k) + \frac{1}{\Lambda(k)} \right]^{-1} \quad (2.3.16)$$

综合 (2.3.12)、(2.3.15)、和 (2.3.16) 式便得到加权最小二乘参数估计递推算法 (简称 RWLS)

$$\begin{cases} \hat{\boldsymbol{\theta}}(k) = \hat{\boldsymbol{\theta}}(k-1) + \mathbf{K}(k)[z(k) - \mathbf{h}^T(k)\hat{\boldsymbol{\theta}}(k-1)] \\ \mathbf{K}(k) = \mathbf{P}(k-1)\mathbf{h}(k) \left[\mathbf{h}^T(k)\mathbf{P}(k-1)\mathbf{h}(k) + \frac{1}{\Lambda(k)} \right]^{-1} \\ \mathbf{P}(k) = [\mathbf{I} - \mathbf{K}(k)\mathbf{h}^T(k)]\mathbf{P}(k-1) \end{cases} \quad (2.3.17)$$

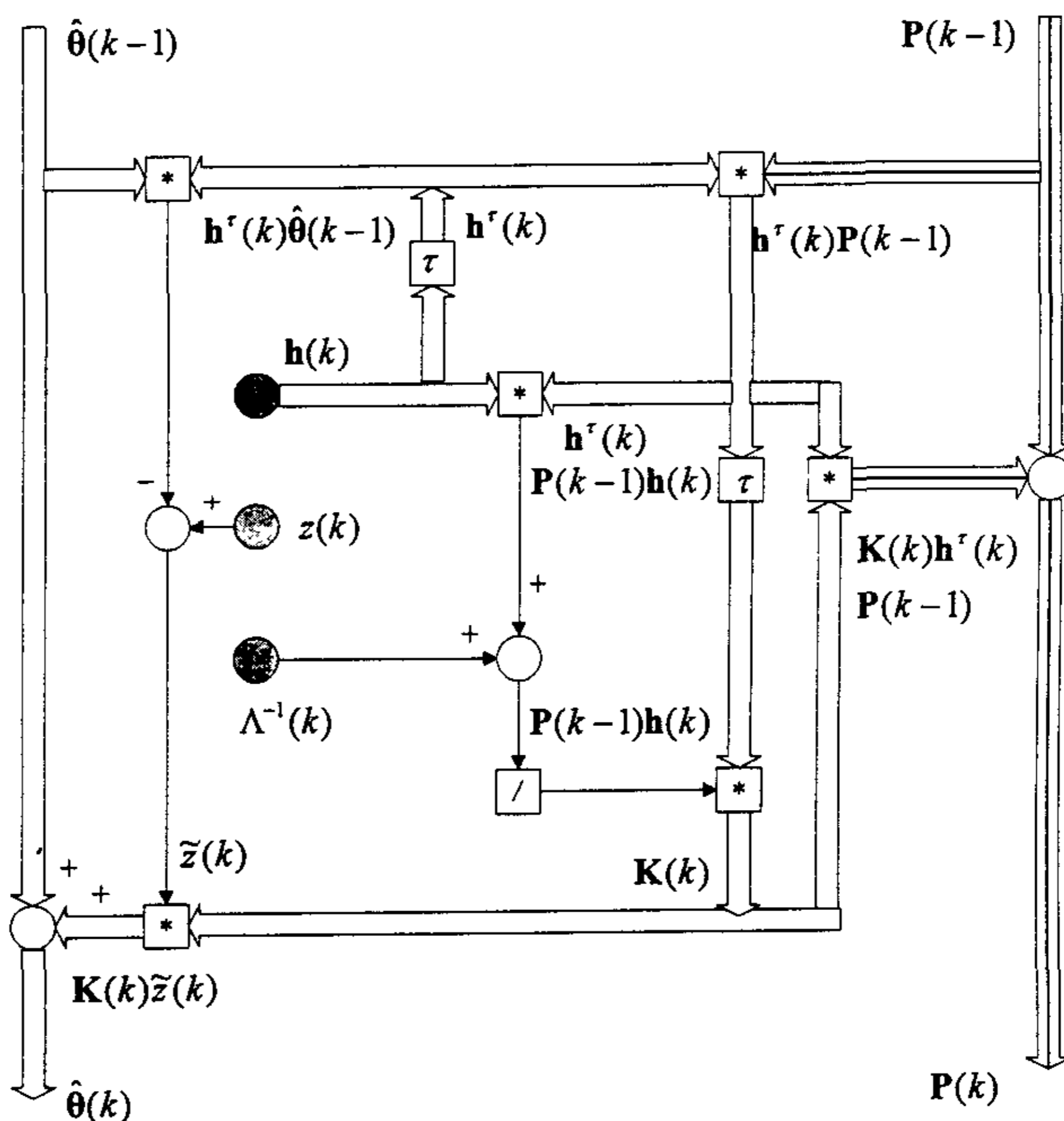
其中, 当 $\Lambda(k) = 1, \forall k$ 时, 加权最小二乘递推算法就退化成最小二乘递推算法 (简称 RLS)。

(2.3.17) 式表明, k 时刻的参数估计值 $\hat{\boldsymbol{\theta}}(k)$ 等于 $(k-1)$ 时刻的参数估计值 $\hat{\boldsymbol{\theta}}(k-1)$ 加上修正项, 修正项正比于 k 时刻的新息 $\tilde{z}(k) \equiv z(k) - \mathbf{h}^T(k)\hat{\boldsymbol{\theta}}(k-1)$, 其增益矩阵 $\mathbf{K}(k)$ 是时变矩阵, $\mathbf{P}(k)$ 是对称矩阵, 为了保证 $\mathbf{P}(k)$ 的对称性, 有时把 (2.3.17) 式中的第 3 算式写成

$$\begin{aligned} \mathbf{P}(k) &= \mathbf{P}(k-1) - \frac{[\mathbf{P}(k-1)\mathbf{h}(k)][\mathbf{P}(k-1)\mathbf{h}(k)]^T}{\mathbf{h}^T(k)\mathbf{P}(k-1)\mathbf{h}(k) + \frac{1}{\Lambda(k)}} \\ &= \mathbf{P}(k-1) - \mathbf{K}(k)\mathbf{K}^T(k) \left[\mathbf{h}^T(k)\mathbf{P}(k-1)\mathbf{h}(k) + \frac{1}{\Lambda(k)} \right] \end{aligned} \quad (2.3.18)$$

这样, 在计算过程中即使有舍入误差, 也能保持 $\mathbf{P}(k)$ 矩阵始终是对称的。

(2.3.17) 是一种递推关系, 其递推计算流程如图 2-1 所示。



图例: * 乘法器 / 除法器 tau 转置 ● 数据
 → 标量信息 ⇒ 向量信息 ⇨ 矩阵信息

图 2-1 递推计算流程

在进行递推计算时，需要事先选择初始状态 $\hat{\theta}(0)$ 和 $P(0)$ ，它们的取值有两种方法。一种是根据一批数据，利用一次完成算法，预先求得

$$\begin{cases} P(L_0) = (H_{L_0}^T \Lambda_{L_0} H_{L_0})^{-1} \\ \hat{\theta}(L_0) = P(0) H_{L_0}^T \Lambda_{L_0} z_{L_0} \end{cases} \quad (2.3.19)$$

置 $P(0) = P(L_0)$ ， $\hat{\theta}(0) = \hat{\theta}(L_0)$ ，其中 L_0 为数据长度，为了减少计算量， L_0 不宜取太大；另一种是直接取

$$\begin{cases} \mathbf{P}(L_0) = a^2 \mathbf{I}, & a \text{ 为充分大的实数} \\ \hat{\boldsymbol{\theta}}(L_0) = \boldsymbol{\varepsilon}, & \boldsymbol{\varepsilon} \text{ 为充分小的实向量} \end{cases} \quad (2.3.20)$$

因为

$$\begin{cases} \mathbf{P}^{-1}(k) = \sum_{i=1}^k \Lambda(i) \mathbf{h}(i) \mathbf{h}^T(k) \\ \mathbf{P}^{-1}(k) \hat{\boldsymbol{\theta}}(k) = \sum_{i=1}^k \Lambda(i) \mathbf{h}(i) z(i) \end{cases} \quad (2.3.21)$$

根据 (2.3.2) 式, 则有

$$\begin{aligned} \hat{\boldsymbol{\theta}}(k) &= \left[\sum_{i=1}^k \Lambda(i) \mathbf{h}(i) \mathbf{h}^T(k) \right]^{-1} \left[\sum_{i=1}^k \Lambda(i) \mathbf{h}(i) z(i) \right] \\ &= \left[\mathbf{P}^{-1}(0) + \sum_{i=1}^k \Lambda(i) \mathbf{h}(i) \mathbf{h}^T(k) \right]^{-1} \left[\mathbf{P}^{-1}(0) \hat{\boldsymbol{\theta}}(0) + \sum_{i=1}^k \Lambda(i) \mathbf{h}(i) z(i) \right] \end{aligned} \quad (2.3.22)$$

显然, 使上式成立的条件是 $\mathbf{P}^{-1}(0) \rightarrow \mathbf{0}$ 及 $\hat{\boldsymbol{\theta}}(0) \rightarrow \mathbf{0}$, 故有 (2.3.20) 式。

另外, 可用下式作为递推算法的停机标准

$$\max_{v_i} \left| \frac{\hat{\theta}_i(k) - \hat{\theta}_i(k-1)}{\hat{\theta}_i(k-1)} \right| < \varepsilon, \quad \varepsilon \text{ 是一个适当小的数} \quad (2.3.23)$$

它意味着当所有的参数估计值变化不大时, 即可停机。

2.3.2 依模型阶次的递推算法

进行模型阶次辨识时, 需要依阶次递推的算法, 即希望在 n 阶模型参数估计的基础上递推计算 $n+1$ 阶模型的参数, 这样能减少计算量。

考虑如下的模型

$$A(z^{-1})z(k) = B(z^{-1})u(k) + v(k) \quad (2.3.24)$$

为简单起见, 设

$$\begin{cases} A(z^{-1}) = 1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_n z^{-n} \\ B(z^{-1}) = b_1 z^{-1} + b_2 z^{-2} + \dots + b_n z^{-n} \end{cases} \quad (2.3.25)$$

且 $v(k)$ 为零均值的白噪声。令

$$\begin{cases} \boldsymbol{\theta}_n = [a_1, b_1, a_2, b_2, \dots, a_n, b_n]^T \\ \boldsymbol{\theta}_{n+1} = [a_1, b_1, a_2, b_2, \dots, a_n, b_n \mid a_{n+1}, b_{n+1}]^T \\ \mathbf{z}_L = [z(1), z(2), \dots, z(L)]^T, L \text{ 为数据长度} \end{cases} \quad (2.3.26)$$

当模型为 n 阶时, 根据 (2.2.10) 式, 可得 n 阶模型参数的最小二乘估计为

$$\hat{\boldsymbol{\theta}}_n = (\mathbf{H}_n^T \mathbf{H}_n)^{-1} \mathbf{H}_n^T \mathbf{z}_L \cong \mathbf{P}_n \mathbf{H}_n^T \mathbf{z}_L \quad (2.3.27)$$

其中

$$\mathbf{H}_n = \begin{bmatrix} -z(0) & u(0) & -z(1) & u(1) & \dots & -z(1-n) & u(1-n) \\ -z(1) & u(1) & -z(2) & u(2) & \dots & -z(2-n) & u(2-n) \\ & & & & \dots & & \\ -z(L-1) & u(L-1) & -z(L-2) & u(L-2) & \dots & -z(L-n) & u(L-n) \end{bmatrix} \quad (2.3.28)$$

当模型为 n 阶时, 根据 (2.2.10) 式, 又有

$$\hat{\boldsymbol{\theta}}_{n+1} = (\mathbf{H}_{n+1}^T \mathbf{H}_{n+1})^{-1} \mathbf{H}_{n+1}^T \mathbf{z}_L \cong \mathbf{P}_{n+1} \mathbf{H}_{n+1}^T \mathbf{z}_L \quad (2.3.29)$$

其中

$$\begin{aligned} \mathbf{H}_{n+1} &= \begin{bmatrix} -z(0) & u(0) & \dots & -z(1-n) & u(1-n) & \vdots & -z(-n) & u(-n) \\ -z(1) & u(1) & \dots & -z(2-n) & u(2-n) & \vdots & -z(1-n) & u(L-n) \\ & & \dots & & & & & \\ -z(L-1) & u(L-1) & \dots & -z(L-n) & u(L-n) & \vdots & -z(L-n-1) & u(L-n-1) \end{bmatrix} \\ &\cong [\mathbf{H}_n \mid \mathbf{X}_n] \end{aligned} \quad (2.3.30)$$

且有

$$\mathbf{P}_{n+1} = (\mathbf{H}_{n+1}^T \mathbf{H}_{n+1})^{-1} = \left(\begin{bmatrix} \mathbf{H}_n^T \\ \mathbf{X}_n^T \end{bmatrix} [\mathbf{H}_n \mid \mathbf{X}_n] \right)^{-1} = \begin{bmatrix} \mathbf{P}_n^{-1} & \vdots & \boldsymbol{\Omega}_n^T \\ \boldsymbol{\Omega}_n & \vdots & \mathbf{Q}_n \end{bmatrix}^{-1} \quad (2.3.31)$$

式中

$$\begin{cases} \mathbf{P}_n = (\mathbf{H}_n^T \mathbf{H}_n)^{-1} \\ \boldsymbol{\Omega}_n = \mathbf{X}_n^T \mathbf{H}_n \\ \mathbf{Q}_n = \mathbf{X}_n^T \mathbf{X}_n \end{cases} \quad (2.3.32)$$

根据分块矩阵求逆公式, 得

$$\mathbf{P}_{n+1} = \begin{bmatrix} \mathbf{P}_n + \mathbf{R}_n \boldsymbol{\Omega}_n \mathbf{P}_n & | & \mathbf{R}_n \\ \hline -\mathbf{R}_n^T & | & \mathbf{A}_n \end{bmatrix} \quad (2.3.33)$$

其中

$$\begin{cases} \mathbf{R}_n = \mathbf{P}_n \boldsymbol{\Omega}_n^T \mathbf{A}_n \\ \mathbf{A}_n = (\mathbf{Q}_n - \boldsymbol{\Omega}_n \mathbf{P}_n \boldsymbol{\Omega}_n^T)^{-1} \end{cases} \quad (2.3.34)$$

根据 (2.3.29) 式, 并注意到 (2.3.33) 式, 得

$$\hat{\boldsymbol{\theta}}_{n+1} = \begin{bmatrix} \mathbf{P}_n + \mathbf{R}_n \boldsymbol{\Omega}_n \mathbf{P}_n & | & \mathbf{R}_n \\ \hline -\mathbf{R}_n^T & | & \mathbf{A}_n \end{bmatrix} \begin{bmatrix} \mathbf{H}_n^T \\ \mathbf{X}_n \end{bmatrix} \mathbf{z}_L \quad (2.3.35)$$

展开之, 利用 (2.3.27) 式, 得

$$\hat{\boldsymbol{\theta}}_{n+1} = \begin{bmatrix} \hat{\boldsymbol{\theta}}_n - \mathbf{R}_n \mathbf{X}_n^T (\mathbf{z}_L - \mathbf{H}_n \hat{\boldsymbol{\theta}}_n) \\ \mathbf{A}_n \mathbf{X}_n^T (\mathbf{z}_L - \mathbf{H}_n \hat{\boldsymbol{\theta}}_n) \end{bmatrix} \quad (2.3.36)$$

(2.3.36) 和 (2.3.33) 式构成了依模型阶次递推的算法。设模型最低的可能阶次为 n^* , 取 $\hat{\boldsymbol{\theta}}_{n^*}(0) = \boldsymbol{\varepsilon}$, $\mathbf{P}_{n^*}(0) = a^2 \mathbf{I}$, 利用依观测次序的递推算法, 递推到 L 时刻, 获得 $\hat{\boldsymbol{\theta}}_{n^*}$ 和 \mathbf{P}_{n^*} ; 再利用 (2.3.36) 和 (2.3.33) 式, 递推计算 $\hat{\boldsymbol{\theta}}_{n^*+1}$ 和 \mathbf{P}_{n^*+1} , 如此继续下去, 便可获得预定阶次的参数估计 $\hat{\boldsymbol{\theta}}_n$, 递推过程中要特别注意数据阵 \mathbf{H}_n 和 \mathbf{X}_n 的构成。

为了减少计算量, 依模型阶次递推算法中的一些矩阵可用相关函数的形式来表示。如果数据是平稳序列, 当数据长度 L 充分大时, 算法中的一些矩阵可用相关函数表示

$$\begin{cases} \mathbf{Q}_n = \mathbf{X}_n^T \mathbf{X}_n = L \begin{bmatrix} R_z(0) & -R_{zu}(0) \\ -R_{zu}(0) & R_u(0) \end{bmatrix} \\ \mathbf{\Omega}_n = \mathbf{X}_n^T \mathbf{H}_n = L \begin{bmatrix} R_z(n) & -R_{zu}(n) & R_z(n-1) & -R_{zu}(n-1) & \dots & R_z(1) & -R_{zu}(1) \\ -R_{uz}(n) & R_u(n) & -R_{uz}(n-1) & R_u(n-1) & \dots & -R_{uz}(1) & R_u(1) \end{bmatrix} \\ \mathbf{X}_n^T \mathbf{z}_L = L \begin{bmatrix} R_z(n+1) \\ R_{uz}(n+1) \end{bmatrix} \end{cases} \quad (2.3.37)$$

上式表明, \mathbf{Q}_n 与阶次 n 无关, 只要先把相关函数 $R_z(0)$ 、 $R_u(0)$ 和 $R_{zu}(0)$ 计

算出来，构造出 Q_n ，则 Q_n 在递推过程中是个常数阵。 $\Omega_n X_n^T$ 和 z_L 虽然与阶次 n 有关，但模型阶次每增加一阶时，只需计算相关函数 $R_z(n+1)$ 、 $R_v(n+1)$ 、 $R_{zw}(n+1)$ 和 $R_w(n+1)$ ，其中

$$\Omega_{n+1} = \begin{bmatrix} LR_z(n+1) & -LR_{zw}(n+1) \\ -LR_{zw}(n+1) & LR_v(n+1) \end{bmatrix} \Omega_n \quad (2.3.38)$$

这样大大减少了计算量。进一步把 (2.3.36) 式写成

$$\hat{\theta}_{n+1} = \begin{bmatrix} \hat{\theta}_n - R_n(X_n^T z_L - \Omega_n \hat{\theta}_n) \\ A_n(X_n^T z_L - \Omega_n \hat{\theta}_n) \end{bmatrix} \quad (2.3.39)$$

它与 (2.3.33) 式构成按相关函数计算的依阶次递推算法。(2.3.39) 和 (2.3.33) 式中的 R_n 和 A_n 仍按 (2.3.34) 式计算，但 Ω_n 和 $X_n^T z_L$ 按 (2.3.37) 式计算， Q_n 为常数阵。

总之，最小二乘法是一种应用非常广泛的辨识方法，无论是一次完成算法，还是递推算法在工程上都是比较容易实现的。此外，最小二乘法还具有一个突出的优点：鲁棒性强。然而，它的缺点也是显著的，如果模型噪声不是白噪声，最小二乘参数估计一般就不再是无偏一致估计。

2.4 其它几种最小二乘类参数辨识方法

最小二乘法是一种最基本的辨识方法，但它具有两方面的缺陷：

(1) 当模型噪声是有色噪声时，最小二乘参数估计不是无偏、一致估计；

(2) 随着数据的增长，最小二乘法将出现所谓的“数据饱和”现象。这是由于增益矩阵 $K(k)$ 随着 k 的增加将逐渐趋近于零，以致递推算法慢慢失去修正能力之故。下面介绍的几种辨识方法，就是针对这两个问题所提出的。它们的基本思想与最小二乘法没有什么根本的区别，但具体做法上

却各有特点。

2.4.1 适应算法

适应算法是针对“数据饱和”现象提出的一种辨识方法，它包括遗忘因子法和限定记忆法两种算法。

所谓“数据饱和”现象就是随着时间的推移，采集到的数据越来越多，新数据所提供的信息被淹没在老数据的海洋之中。如果辨识算法对新、老数据给予相同的信度，那么随着从新数据中获得的信息量相对下降，算法慢慢就会失去修正能力。最小二乘法就是这样对待新、老数据的，因此在实际应用中常常会出现“数据饱和”现象。这是参数估计值可能还偏离真值较远就无法更新了；对时变过程来说，它又将导致参数估计值不能跟踪时变参数的变化。

1、遗忘因子法

这种方法的基本思想是对老数据加上遗忘因子，以降低老数据所提供的信息量，增加新数据的信息量。这里仅介绍遗忘因子的递推算法(RFF):

$$\begin{cases} \hat{\theta}(k) = \hat{\theta}(k-1) + \mathbf{K}(k)[z(k) - \mathbf{h}^T(k)\hat{\theta}(k-1)] \\ \mathbf{K}(k) = \mathbf{P}(k-1)\mathbf{h}(k)[\mathbf{h}^T(k)\mathbf{P}(k-1)\mathbf{h}(k) + \mu]^{-1} \\ \mathbf{P}(k) = \frac{1}{\mu}[\mathbf{I} - \mathbf{K}(k)\mathbf{h}^T(k)]\mathbf{P}(k-1) \\ 0 < \mu \leq 1 \end{cases} \quad (2.4.1)$$

式中的 μ 即为遗忘因子。

2、限定记忆法

最小二乘法或遗忘因子法在一次完成算法中所用的数据长度 L 是一定的。但在递推算法中，数据长度 L 就不是固定的了，它随着时间 k 的推移而逐渐增加。这意味着老数据所含的信息在不断累积，长期下去新数据所含的信息将被淹没，新数据的作用就会被削弱。这种数据长度随 k 不断增长的辨识算法称作增长记忆法，其特点是老数据所含的信息始终在起作

用，相对将影响新数据的作用。另一类辨识算法叫作限定记忆法。这种方法参数估计值仅仅依赖于有限个最新数据所提供的信息，每增加一个新数据的信息，就要去掉一个老数据的信息，数据长度始终保持不变。它的特点是离现时刻 L 以前的老数据所含的信息从算法中彻底被抛弃，影响参数估计值的数据始终是最新的 L 个数据，不像最小二乘法或遗忘因子法那样，不管多老的数据都在起作用。就这点而言，限定记忆法更适合于用来克服“数据饱和”现象。

其递推算法 (RFM) 如下：

$$\begin{cases}
 \hat{\theta}(k+1, k+L) = \hat{\theta}(k, k+L) - \mathbf{K}(k+1, k+L)[z(k) - \mathbf{h}^T(k)\hat{\theta}(k, k+L)] \\
 \mathbf{K}(k+1, k+L) = \mathbf{P}(k, k+L)\mathbf{h}(k)[1 - \mathbf{h}^T(k)\mathbf{P}(k, k+L)\mathbf{h}(k)]^{-1} \\
 \mathbf{P}(k+1, k+L) = [\mathbf{I} + \mathbf{K}(k+1, k+L)\mathbf{h}^T(k)]\mathbf{P}(k, k+L) \\
 \hat{\theta}(k, k+L) = \hat{\theta}(k, k+L-1) - \mathbf{K}(k, k+L)[z(k+L) - \mathbf{h}^T(k+L)\hat{\theta}(k, k+L-1)] \\
 \mathbf{K}(k, k+L) = \mathbf{P}(k, k+L-1)\mathbf{h}(k+L)[1 + \mathbf{h}^T(k+L)\mathbf{P}(k, k+L-1)\mathbf{h}(k+L)]^{-1} \\
 \mathbf{P}(k, k+L) = [\mathbf{I} - \mathbf{K}(k, k+L)\mathbf{h}^T(k+L)]\mathbf{P}(k, k+L-1)
 \end{cases}$$

(2.4.2)

2.4.2 偏差补偿最小二乘法

本小节和后面介绍的几种算法，都是针对最小二乘法就有色噪声不能给出无偏一致估计而提出的最小二乘类参数辨识方法。

偏差补偿最小二乘法是针对过程存在均值为零，方差为 σ_w^2 的不相关随机测量噪声的。其递推算法 (RCLS) 可归纳成

$$\left\{ \begin{aligned}
 \hat{\theta}_{LS}(k) &= \hat{\theta}_{LS}(k-1) + \mathbf{K}(k)[z(k) - \mathbf{h}^T(k)\hat{\theta}_{LS}(k-1)] \\
 \mathbf{K}(k) &= \mathbf{P}(k-1)\mathbf{h}(k)[1 + \mathbf{h}^T(k)\mathbf{P}(k-1)\mathbf{h}(k)]^{-1} \\
 \mathbf{P}(k) &= [\mathbf{I} - \mathbf{K}(k)\mathbf{h}^T(k)]\mathbf{P}(k-1) \\
 J(k) &= J(k-1) + \frac{[z(k) - \mathbf{h}^T(k)\hat{\theta}_{LS}(k-1)]^2}{1 + \mathbf{h}^T(k)\mathbf{P}(k-1)\mathbf{h}(k)} \\
 \hat{\sigma}_w^2(k) &= \frac{J(k)}{k[1 + \hat{\theta}_C^T(k)\mathbf{D}\hat{\theta}_{LS}(k)]} \\
 \mathbf{D} &= \begin{bmatrix} \mathbf{I}_{n_a} & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{n_b} \end{bmatrix} \\
 \hat{\theta}_C(k) &= \hat{\theta}_{LS}(k) + k\hat{\sigma}_w^2(k)\mathbf{P}(k)\mathbf{D}\hat{\theta}_C(k-1)
 \end{aligned} \right. \quad (2.4.3)$$

2.4.3 增广最小二乘法

如果 SISO 过程采用如下的数学模型

$$A(z^{-1})z(k) = B(z^{-1})u(k) + D(z^{-1})v(k) \quad (2.4.4)$$

描述，其中 $u(k)$ 和 $z(k)$ 表示过程的输入和输出； $v(k)$ 是均值为零的不相关随机噪声；且

$$\left\{ \begin{aligned}
 A(z^{-1}) &= 1 + a_1z^{-1} + a_2z^{-2} + \dots + a_{n_a}z^{-n_a} \\
 B(z^{-1}) &= b_1z^{-1} + b_2z^{-2} + \dots + b_{n_b}z^{-n_b} \\
 D(z^{-1}) &= 1 + d_1z^{-1} + d_2z^{-2} + \dots + d_{n_d}z^{-n_d}
 \end{aligned} \right. \quad (2.4.5)$$

并且假定模型的阶次 n_a 、 n_b 和 n_d 已经确定，则这类问题的辨识可用增广最小二乘法。其递推算法 (RELS) 如下

$$\left\{ \begin{aligned}
 \hat{\theta}(k) &= \hat{\theta}(k-1) + \mathbf{K}(k)[z(k) - \mathbf{h}^T(k)\hat{\theta}(k-1)] \\
 \mathbf{K}(k) &= \mathbf{P}(k-1)\mathbf{h}(k)[\mathbf{h}^T(k)\mathbf{P}(k-1)\mathbf{h}(k) + 1]^{-1} \\
 \mathbf{P}(k) &= [\mathbf{I} - \mathbf{K}(k)\mathbf{h}^T(k)]\mathbf{P}(k-1)
 \end{aligned} \right. \quad (2.4.6)$$

这是最小二乘法的一种简单推广。

2.4.4 广义最小二乘法

设 SISO 过程采用如下的数学模型

$$A(z^{-1})z(k) = B(z^{-1})u(k) + \frac{1}{C(z^{-1})}v(k) \quad (2.4.7)$$

描述, 其中 $u(k)$ 和 $z(k)$ 表示过程的输入和输出; $v(k)$ 是均值为零的不相关随机噪声; 且

$$\begin{cases} A(z^{-1}) = 1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_{n_a} z^{-n_a} \\ B(z^{-1}) = b_1 z^{-1} + b_2 z^{-2} + \dots + b_{n_b} z^{-n_b} \\ C(z^{-1}) = 1 + c_1 z^{-1} + c_2 z^{-2} + \dots + c_{n_c} z^{-n_c} \end{cases} \quad (2.4.8)$$

并且假定模型的阶次 n_a 、 n_b 和 n_c 已经确定, 则这类问题的辨识可用广义最小二乘法, 以便获得无偏一致估计。广义最小二乘递推算法 (RGLS) 可归纳成

$$\begin{cases} \hat{\theta}(k) = \hat{\theta}(k-1) + \mathbf{K}_f(k)[z_f(k) - \mathbf{h}_f^T(k)\hat{\theta}(k-1)] \\ \mathbf{K}_f(k) = \mathbf{P}_f(k-1)\mathbf{h}_f(k)[1 + \mathbf{h}_f^T(k)\mathbf{P}_f(k-1)\mathbf{h}_f(k)]^{-1} \\ \mathbf{P}_f(k) = [\mathbf{I} - \mathbf{K}_f(k)\mathbf{h}_f^T(k)]\mathbf{P}_f(k-1) \\ \hat{\theta}_e(k) = \hat{\theta}_e(k-1) + \mathbf{K}_e(k)[\hat{e}(k) - \mathbf{h}_e^T(k)\hat{\theta}_e(k-1)] \\ \mathbf{K}_e(k) = \mathbf{P}_e(k-1)\mathbf{h}_e(k)[1 + \mathbf{h}_e^T(k)\mathbf{P}_e(k-1)\mathbf{h}_e(k)]^{-1} \\ \mathbf{P}_e(k) = [\mathbf{I} - \mathbf{K}_e(k)\mathbf{h}_e^T(k)]\mathbf{P}_e(k-1) \\ \hat{e}(k) = \frac{1}{C(z^{-1})}v(k) \end{cases} \quad (2.4.9)$$

2.5 本章小结

本章讨论了最小二乘类参数辨识方法, 除了上面讲的几种, 还有辅助变量法、二步法、多级最小二乘法、Yule-Walker 辨识算法。这些方法的递推算法都有共同的结构: 新的参数估计值=老的参数估计值+增益矩阵×新息。下面将几种主要的辨识方法的性能比较如表 2-2 所示。

表 2-2 几种最小二乘类辨识方法性能比较

辨识方法	性能
LS	①如果模型噪声是白噪声, 则 $\hat{\theta}_{LS} \xrightarrow{a.s.} \theta_0$
	②如果模型噪声是有色噪声, 则 $\hat{\theta}_{LS}$ 是有偏估计量
	③鲁棒性强, 易实现
	④计算量较小
	⑤对数据直流分量敏感
	⑥对高阶模型来说, 性能明显优于其他方法, 而且有较可靠的收敛性
	⑦数据无需过多, 过多的数据并不改善辨识精度
	⑧常用作其他辨识方法的起步, 以获得他种方法的初始值
	①当噪信比较小时, 辨识精度高
	②当噪信比较大时, 收敛点可能不唯一, 参数估计值往往是有偏的
GLS	③计算量比较大
	④数据要充分多, 否则辨识精度明显下降
	⑤初始值 $P_f(0)$ 和 $P_e(0)$ 对辨识结果有较大影响
	⑥噪声模型阶次不宜过高
ELS	①一般情况下, $\hat{\theta} \xrightarrow{a.s.} \theta_0$, 但噪声模型参数不能估计很准
	②多数性能与 LS 类似
	③ELS 是极大似然法的一种近似形式。当数据长度较小时, 辨识精度可能优于极大似然法; 但数据长度较大时, 精度低于极大似然法

3 多输入多输出系统辨识方法

本章在前面的基础上，提出了一种全新的多输入多输出系统辨识方法。目前的辨识方法一般需要在系统输入端加入激励信号，而且针对多输入多输出系统的在线辨识仍很困难。本文提出一种基于牛顿迭代法的多输入、多输出对象模型迭代辨识方法，模型参数更新的依据是使模型预测输出与全部采样时刻的对象实际输出之间的均方差递减，直到收敛。这种基于全局数据迭代的辨识方法可进行闭环辨识，无需外加激励信号，适用于多输入多输出对象的在线辨识。其优点还在于可以充分利用一段数据，而不用考虑数据长度对辨识精度的影响的问题。对一个两输入、两输出对象模型的仿真研究的计算结果表明，辨识效果令人满意。

3.1 问题的提出

在火力发电厂中，从燃料、供风和给水等进入锅炉到机组输出电力的过程是一个典型的多输入、多输出、强耦合的复杂过程，其优化控制的研究一直是热工自动化工作者几十年来关注的焦点。

热工过程自动控制的发展，经历了以下阶段^[22]：

(1) 初级阶段：本世纪 50 年代后前后，热工生产过程主要是凭生产实践经验，局限于一般的控制零件及机电式控制仪器，采用比较笨重的机电式仪表实现机、炉、电各自独立的分散的局部自动控制。机、炉、电各控制系统之间没有或很少有联系。过程控制的目的是几种热工参数，如温度、压力、流量及液位的定值控制，以保证产品质量和产量的稳定，所应用的理论为古典控制理论。

(2) 仪表化阶段：50 年代末及以后十年间，先后出现了电动单元组合仪表和巡回检测装置，因而实现了把机、炉作为一单元整体来进行集中

控制，仪表盘装在一起监视，从而使机、炉启停运行更为协调，对提高设备效率和强化生产过程有所促进，适应了工业生产设备日益大型化与连续化发展的需要。随着仪表工业的迅速发展对过程控制及对象特性的认识，对仪表及控制系统的设计计算方法都有了较快的发展，而且随着机组拉容量的增大，集中控制机、炉又进一步发展为机、炉、电集中控制。此时所用的仪表有电动及组装仪表，所应用的理论为古典控制理论。

(3) 综合自动化阶段：70年代至今，由于集成电路及计算机技术的飞速发展，由分散的机组或车间控制，向全车间甚至全企业的综合自动化发展，实现了过程控制最优化与管理调度自动化相结合的分散计算机控制。这是过程控制发展的一个新阶段。对电厂而言，则是把火电厂的生产过程（包括主、辅机，全厂各辅助车间）作为一个整体来进行控制，此时所用的仪表有气动仪表、电动组合仪表、组装仪表及计算机，所用理论大多为古典控制理论，少量为现代控制理论。

随着计算机技术的迅速发展，电厂热工过程控制又经历了以下几个计算机控制过程：

集中型计算机控制：它是用一台计算机对整个生产过程进行整体控制，因此对计算机的可靠性要求很高，一旦计算机出现事故，将使整个生产受到影响。

分散型计算机控制：随着微机的大批生产，成本的不断降低，逐渐把集中控制改为用微机进行局部控制，克服了集中控制的一些缺点，但此时各系统之间很难协调起来。

计算机分散控制：它把各系统之间、厂级管理、调度等用一台功能很强的计算机进行上位管理；而把各子系统用微机控制，充分发挥了集中控制和分散控制各自的优点，是一种比较合理的控制方法。

以自适应控制为代表的大量新型控制系统在过程控制中成功应用的关键问题是快速、实时在线辨识对象模型。开环测试显然已不适应实时控

制与高性能指标控制的要求，尤其是强迫性加入输入信号会造成人为扰动，影响正常的生产过程。最小二乘法在模型辨识中得到了广泛应用，但目前的研究一般针对单输入、单输出系统。文献[21]针对作为单输入、单输出系统的汽轮机调节系统的模型辨识问题，提出在单机运行时利用负荷调整信号、并列运行时利用同步器给定扰动信号作为系统的激励信号，采用相关辨识方法获取对象的模型参数。由于火电机组时时刻刻均处于不断的调整过程中，文献[21]的意义可能还在于：无需外界激励信号，直接利用处于不断调整中的系统闭环运行条件下的输入、输出信号，可进行有效的模型辨识。本文提出一种基于全局数据迭代的模型辨识方法，应用于双输入、双输出系统，可以很好地解决多输入多输出系统内部存在的耦合问题。不需要加入激励信号，从而避免了输入外部激励信号对系统正常工作的影响。只要数据序列含有辨识所需要的足够信息量，就可以依据预测输出与实际输出之间的偏差，应用牛顿迭代法进行迭代，直到由得到模型的预测输出与实际输出之间十分接近为止。

以一个多输入多输出系统为研究对象，其输入为 u_1, u_2, \dots, u_N ，输出为 y_1, y_2, \dots, y_M 。则对象包含 $N \times M$ 个传递函数，设 u_i 到 y_j 的 n 阶传递函数为：

$$G_{i,j}(s) = \frac{b_0 + b_1s + b_2s^2 + \dots + b_{n-1}s^{n-1}}{1 + a_0s + a_1s^2 + a_2s^3 + \dots + a_{n-1}s^n} \quad (3.1.1)$$

这样就可以得到一个表征对象动态特性的模型参数的列向量 X ，它包含 $2n \times (N \times M)$ 个元素：

$$X: [(b_0, b_1, \dots, b_{n-1}, a_0, a_1, \dots, a_{n-1})_{11}, \dots, (b_0, b_1, \dots, b_{n-1}, a_0, a_1, \dots, a_{n-1})_{ij}, \dots, (b_0, b_1, \dots, b_{n-1}, a_0, a_1, \dots, a_{n-1})_{N \times M}]^T \quad (3.1.2)$$

在现场进行采样，采样周期为 ΔT ，采样次数为 N_0 ，并将所有输入输出向量分别迭加，则可以构成如 (3.1.3) 式所示的输入输出列向量，输入向量 U ，维数为 $(N \times N_0) \times 1$ ，输出向量 Y ，维数为 $(M \times N_0) \times 1$ ，记录输

入输出数据序列存于这两个向量中:

$$\begin{aligned} \mathbf{U} &= \{u_i, i=1, \dots, N \times N_0\} = [u_{1,1}, \dots, u_{1,N_0}, u_{2,1}, \dots, u_{2,N_0}, \dots, u_{N,1}, \dots, u_{N,N_0}]^T \\ \mathbf{Y} &= \{y_j, j=1, \dots, M \times N_0\} = [y_{1,1}, \dots, y_{1,N_0}, y_{2,1}, \dots, y_{2,N_0}, \dots, y_{M,1}, \dots, y_{M,N_0}]^T \end{aligned} \quad (3.1.3)$$

这样辨识问题就转化为, 根据输入输出列向量, 求式 (3.1.2) 所表示的 $2n \times N \times M$ 个系数的问题。递推最小二乘辨识方法的特点是按时间顺序来取出成段的数据进行迭代更新对象模型参数[23], 这样在时间上靠后的数据对模型参数的影响就比靠前的数据要大。数据序列一旦用完, 辨识也告结束。本文所提出的基于全局数据迭代的辨识方法, 在给定一个迭代初始模型后, 每次迭代计算中都用到全部数据序列, 从而充分发掘了采样数据反映对象动态特性的潜力。

3.2 辨识方法

我们在计算之前, 要对输入数据进行如下处理

$$\mathbf{U}_i = \mathbf{U}_i - \bar{\mathbf{U}}_i, \quad i=1, 2, \dots, N \quad (3.2.1)$$

式中, $\bar{\mathbf{U}}_i = \frac{\sum_{j=1}^{N_0} u_{i,j}}{N_0}$ 。这样做的目的是, 在负荷波动较小的情况下, 我们要辨识的是系统的动态特性, 所以用 (3.2.1) 式处理后就可以顺利地辨识出系统的动态特性。

为了能够在计算机中处理数据, 必须把拉氏传函进行 Z 变换。其具体方法为

$$s = \frac{2}{\Delta T} \times \frac{1-z^{-1}}{1+z^{-1}} \quad (3.2.2)$$

式中 ΔT 为采样时间。

则 (3.1.1) 式变成

$$\frac{y_j(z^{-1})}{u_i(z^{-1})} = \frac{\frac{b_0 + 2b_1/\Delta t}{m} + \frac{2b_0}{m}z^{-1} + \frac{b_0 - 2b_1/\Delta t}{m}z^{-2} + \dots}{1 + \frac{2 - 8a_1/\Delta t^2}{m}z^{-1} + \frac{1 - 2a_0/\Delta t + 4a_1/\Delta t^2}{m}z^{-2} + \dots} \quad (3.2.3)$$

式中, m 为归一化系数, 它有如下形式:

$$m = 1 + \frac{2a_0}{\Delta T} + \frac{4a_1}{\Delta T^2} + \frac{8a_2}{\Delta T^3} + \dots + \frac{2^n a_{n-1}}{\Delta T^n} \quad (3.2.4)$$

根据牛顿迭代法的思想, 给定一个初始的 \mathbf{X} 向量, 在迭代进行到第 k 步后, 模型参数为 $\mathbf{X}^{(k)}$, 需计算模型的更新量 $\Delta\mathbf{X}^{(k+1)}$, 以获得新的模型参数 $\mathbf{X}^{(k+1)}$:

$$\mathbf{X}^{(k+1)} = \mathbf{X}^{(k)} + \Delta\mathbf{X}^{(k+1)} \quad (3.2.5)$$

模型的更新量 $\Delta\mathbf{X}^{(k+1)}$ 满足

$$\mathbf{Y} \approx \mathbf{Y}'^{(k)} + \frac{\partial \mathbf{Y}'^{(k)}}{\partial \mathbf{X}_i} \Delta\mathbf{X}^{(k+1)} \quad (3.2.6)$$

式中, $\mathbf{Y}'^{(k)}$ 为根据 $\mathbf{X}^{(k)}$ 和实际输入向量 \mathbf{U} 得到的预测输出向量, 即 $\mathbf{Y}'^{(k)} = f(\mathbf{X}^{(k)}, \mathbf{U})$, 迭代的目标, 就是要让 (3.2.6) 式的右边不断地逼近左边。

(3.2.6) 式可以写成

$$\Delta\mathbf{Y}^{(k)} = \mathbf{A}\Delta\mathbf{X}^{(k+1)} \quad (3.2.7)$$

其中 $\Delta\mathbf{Y}^{(k)} = \mathbf{Y} - \mathbf{Y}'^{(k)}$, \mathbf{A} 为敏感性矩阵: $a_{ji} = \frac{\partial Y_j^{(k)}}{\partial X_i}$, $\mathbf{A} \in R^{M \times N_0, 2n \times N \times M}$ 。

为了求出 \mathbf{A} , 作如下考虑: 由于 \mathbf{A} 的元素都是偏导数, 所以一次只将向量 \mathbf{X} 中的一个元素变化一个微小量 (例如 5%), 其它的元素都不变, 此

即为 δx_i ，并由此得到输出向量 Y'' ，再计算出 Y'' 与第 k 步预测输出 $Y^{(k)}$ 之间的差，这就是 A 中的 $\delta Y^{(k)}$ ，即如 (3.2.8) 式所示：

$$\frac{\partial Y^{(k)}}{\partial x_i} \approx \frac{\delta Y^{(k)}}{\delta x_i} = \frac{Y'' - Y^{(k)}}{\delta x_i} \quad (3.2.8)$$

这样就可以计算出 A 矩阵的一列。再回到开始，保持其它元素不变，而只变化 X 中的下一个元素，重复这个过程，可以把 A 矩阵求出。由最小二乘法的原理知道：

$$\Delta X^{(k+1)} = (A^T A)^{-1} A^T \Delta Y^{(k)} \quad (3.2.9)$$

当求出了 $\Delta X^{(k+1)}$ ，并根据 (3.2.5) 式，就可以得到一个新的传函系数向量 $X^{(k+1)}$ 。

由于在开始迭代的时候给定的初始模型与真实模型会有较大的偏差，而我们是采用线性方法来解决非线性问题，所以从起始点开始迭代时，会产生修正幅度过大的问题。因而要逼近模型的最优值，就必须从起始点向最优值方向以较小幅度进行迭代。所以用 (3.2.10) 式对每次迭代的修正幅度进行控制：

$$\Delta X^{(k+1)} = \begin{cases} \Delta X^{(k+1)} & \left(\frac{\sum \Delta x_i^{(k+1)}}{\sum x_i^{(k)}} \right) \leq 0.05 \\ \Delta X^{(k+1)} \times 0.05 / \left(\frac{\sum \Delta x_i^{(k+1)}}{\sum x_i^{(k)}} \right) & \left(\frac{\sum \Delta x_i^{(k+1)}}{\sum x_i^{(k)}} \right) > 0.05 \end{cases} \quad (3.2.10)$$

迭代的结束条件，借鉴最小二乘法的收敛条件，即当 (3.2.11) 式成立时，结束迭代，其中 p 可以取一个较小值，例如 0.001。但这只是一个收敛条件，后面的计算证明了，(3.2.12) 式的收敛条件更经常地被用到，这时也认为已经迭代成功，并结束计算，得到了最优解：

$$\left(\frac{\sum \Delta x_i^{(k+1)}}{\sum \Delta x_i^{(k)}} \right) \leq p \quad (3.2.11)$$

$$\|Y - Y^{(k+1)}\| \geq \|Y - Y^{(k)}\| \quad (3.2.12)$$

至此，所有问题都已经解决。我们可以只利用实际测得的多输入多输出数据，就能够进行全局数据迭代，从而辨识出精确的对象模型。

总结全局数据迭代辨识方法如下：

- (1) 记录对象的实际输入输出序列 U 和 Y ，设置模型初始值 $X^{(0)}$ （式 (3.1.5)、(3.1.6)）；
- (2) 由 $X^{(k)}$ 得到预测输出 $Y^{(k)}$ ，计算出 $\Delta Y^{(k)}$ ；
- (3) 计算矩阵 A （式 (3.2.8)）；
- (4) 根据 $\Delta Y^{(k)}$ 和 A 计算出 $\Delta X^{(k+1)}$ ，如有必要，则对 $\Delta X^{(k+1)}$ 的大小进行限制（式 (3.2.9)、(3.2.10)）；
- (5) 得到 $X^{(k+1)}$ （式 (3.2.5)）；
- (6) 只要满足 (3.2.11)、(3.2.12) 式中的一个，即结束计算，得到最优解 $X^{(k+1)}$ ；如果两式都不满足，则返回第 2 步继续计算。

其计算流程图如下

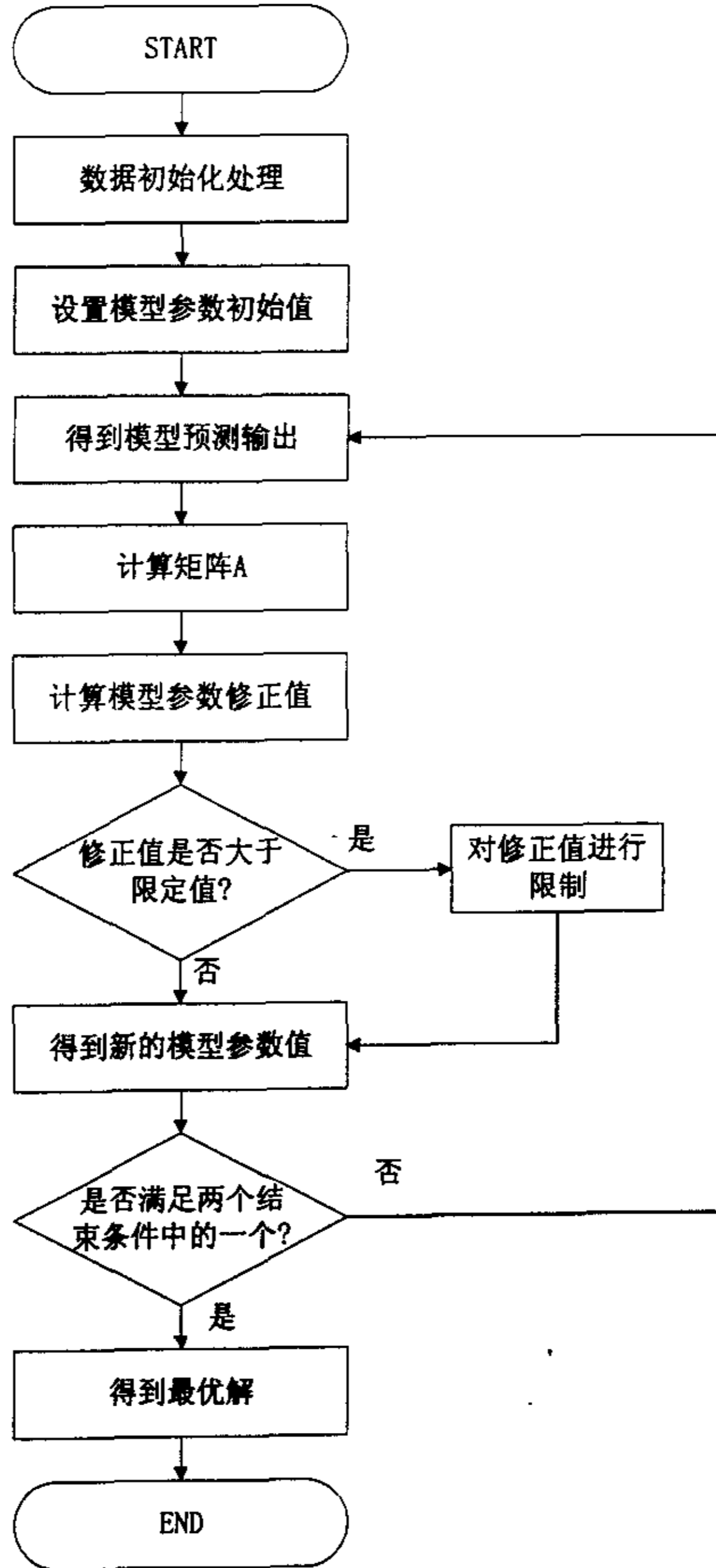


图 3-1 多输入多输出计算方法流程图

3.3 已知模型研究

在如图 3-2 所示的一个双输入双输出系统中，先给出一个理想的输入条件，即两输入分别为正弦和余弦信号。为了使计算结果更有可读性，对输入信号如下处理：

$$\begin{aligned} U_1 &= 10 \sin(0.1i) \\ U_2 &= 10 \cos(0.1i) \end{aligned} \tag{3.3.1}$$

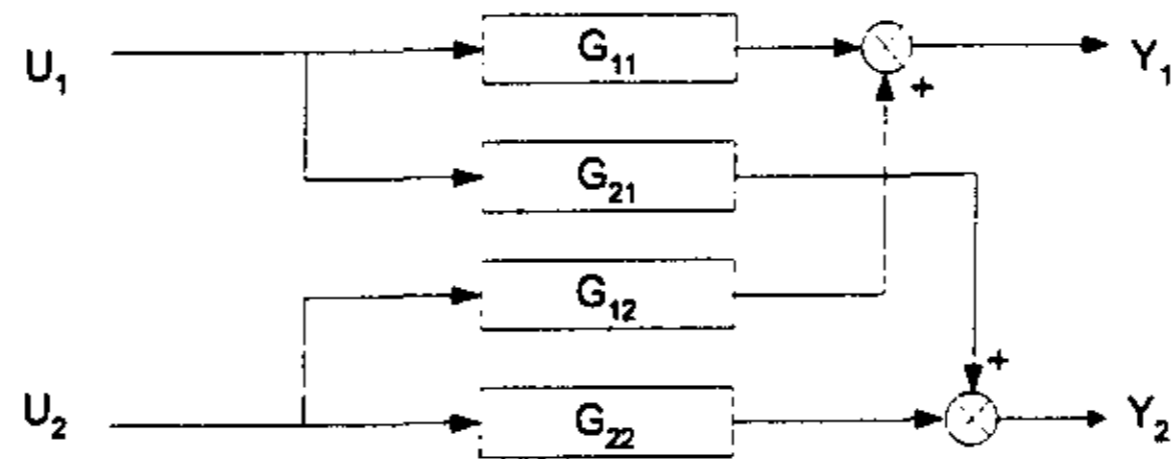


图 3-2 双输入双输出系统模型

我们采用如下的模型结构：

$$G_{i,j}(s) = \frac{b_0 + b_1s}{1 + a_0s + a_1s^2} \tag{3.3.2}$$

给定一个模型，并设之为“实际”对象的模型：

$X: \{\{3, 10, 70, 1200\}, \{-0.2, -5, 90, 430\}, \{12, 1, 200, 5500\}, \{0.9, 1, 110, 8000\}\}$ 。

有了这些我们就可以得到“实际”的输出 Y_1 和 Y_2 。我们把 U_1 、 U_2 、 Y_1 、 Y_2 代入全局数据迭代辨识程序中，并设定一个与“实际”模型有偏差的迭代初始模型：

$X: \{\{4, 7, 100, 1000\}, \{-0.4, -7, 140, 700\}, \{8, 2, 100, 8000\}, \{1.3, 0.6, 200, 5000\}\}$ 可以看出，初始模型与“实际”模型有很大的差别。在这种恶劣的条件下，计算共迭代了 81 步，用了 2 分 27 秒，得到的模拟输出与实际输出之间的均方差为 1.62，如图 3-3 所示，结果拟合的很好。

我们再把理想数据加上 10% 以内的随机误差，得到的拟合均方差为 3.7032，如图 3-4 所示，结果也是非常好。

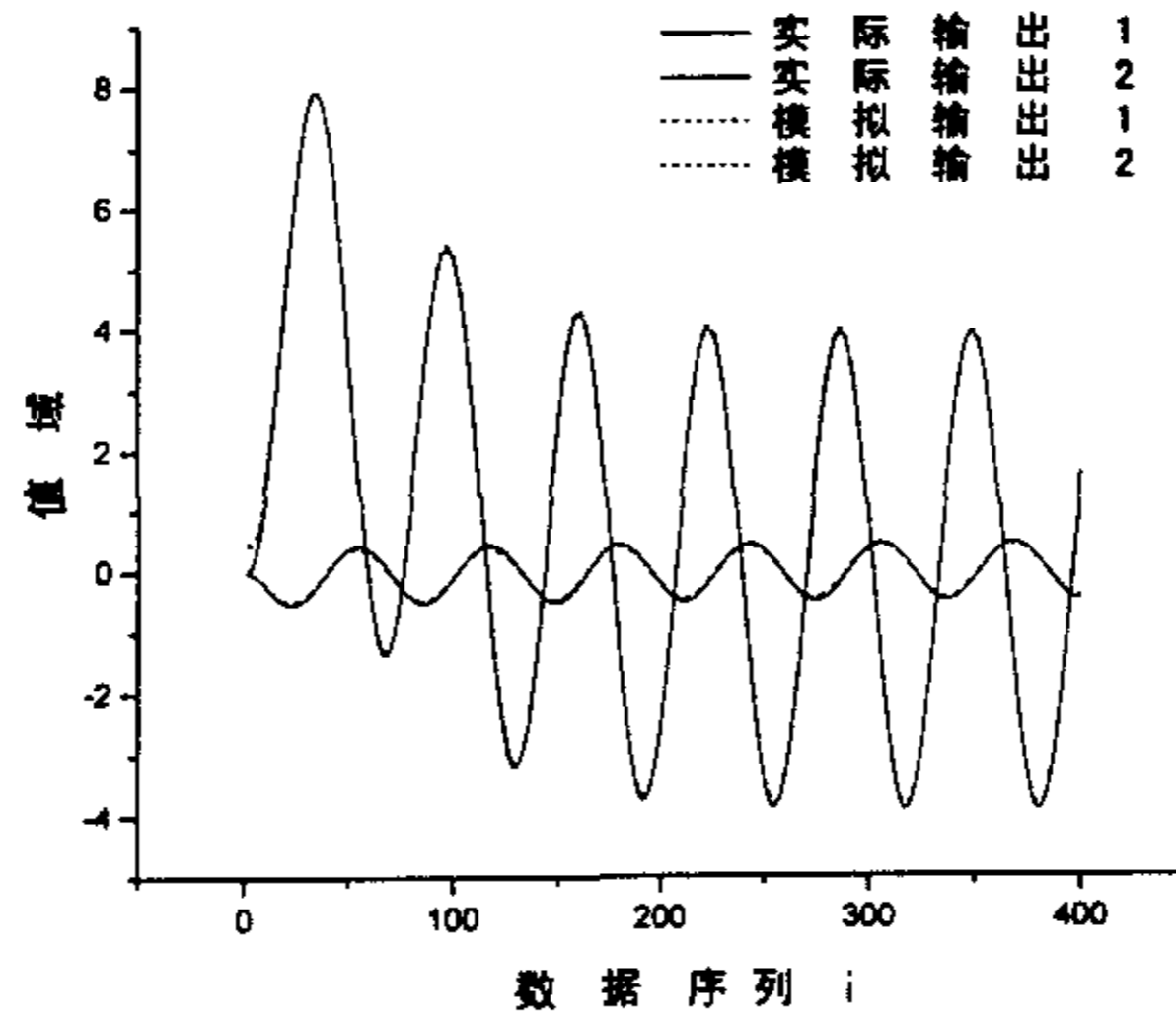


图 3-3、已知模型下的实际输出与模拟输出的比较，均方差 1.6200

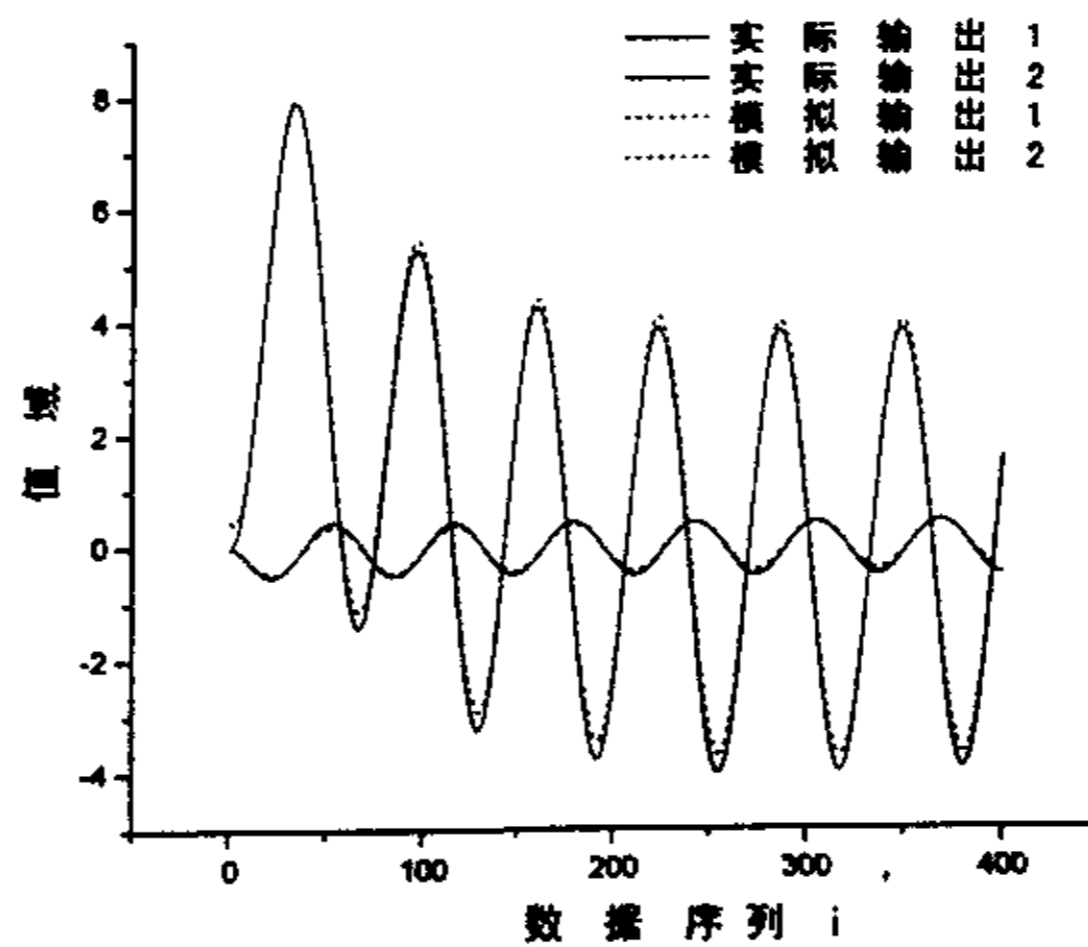


图 3-4、加 10%误差的已知模型下的实际输出与模拟输出的比较，均方差 3.7032

3.4 本章小结

本章在前面论述的最小二乘法的基础上，提出了一种基于牛顿迭代法的全新的多输入多输出系统辨识方法。目前的辨识方法一般需要在系统输入端加入激励信号，而且针对多输入多输出系统的在线辨识仍很困难。这种多输入、多输出对象模型迭代辨识方法，模型参数更新的依据是使模型

预测输出与全部采样时刻的对象实际输出之间的均方差递减，直到收敛。这种基于全局数据迭代的辨识方法可进行闭环辨识，无需外加激励信号，适用于多输入多输出对象的在线辨识。其优点还在于可以充分利用一段数据，而不用考虑数据长度对辨识精度的影响的问题。它可以充分挖掘一段固定长度的数据反应模型参数的能力。并且对一个两输入、两输出对象模型进行了仿真研究，其计算结果表明，辨识效果令人满意。

4 在火电机组双输入双输出过程中的应用

本文在前面一章提出了一种新的基于牛顿迭代法的多输入多输出系统辨识方法，它与其它辨识方法相比，具有可以针对多输入多输出系统进行闭环辨识、无需外加激励信号、无须考虑数据长度对辨识精度的影响等优点。在本章中，将提出一种采用辐射能反馈信号的火电机组负荷控制系统，并基于这种思想，把多输入多输出系统辨识方法应用到火电机组双输入双输出过程中去，计算结果令人满意。

4.1 采用辐射能反馈信号火电机组负荷控制系统

目前作为火电站最重要最关键技术设备的燃煤锅炉，仍然普遍存在着运行效率低、污染指数高（污染物排放量大），缺少优化组织的燃烧控制系统，能源利用率远远低于发达国家的情况，并且已经造成了极其严重的环境污染。从锅炉燃料进料，空气送入，燃料燃烧到热能交换，产生蒸汽，直到以预期的性能指标供应负荷是一个相当复杂的过程，很难用一个特定的对象模型精确描述。要保证锅炉安全、高效地运行，就必须有一种快速、高品质的控制系统。如果对之进行简化，我们不难发现其中包含着具有明显热力系统特征的大时间常数惯性环节、大迟延时间迟延环节和积分环节等等。通常，锅炉入炉燃料是通过汽压的变化来控制的。一旦燃料量发生变化，或者由于燃料品质变化造成燃烧扰动，通过测量锅炉蒸汽参数变化再来调节燃料配给，会带来很大的迟延，造成控制的失调，使锅炉的正常运行产生一系列问题。长时间的入炉燃料量与锅炉负荷的不协调可能使锅炉运行参数难以达到使用要求。多余的燃料进入炉膛的结果会引起燃烧温度升高，给炉管带来超温和超压的危害，使用寿命降低，严重时导致爆管事故的发生，不仅使燃烧过程的安全性失去保障，也将会降低锅炉运行的

经济性。分析锅炉燃烧到蒸汽出力的过程，最能直接地反映入炉燃料综合状况的应该是燃料在炉内燃烧释放能量的过程，这是燃料对锅炉燃烧系统的影响和作用的过程。炉内辐射能是燃烧过程中的一项重要综合参数，它不仅能够快速地反映入炉燃料量的瞬态变化，也能及时地反映由于燃料品质的突变造成的燃烧率瞬态变化。

针对上述情况，华中科技大学周怀春教授提出了一种采用辐射能信号反馈信号建立起来的火电机组负荷控制系统：直接提取炉内辐射能变化，并将炉膛火焰携带的辐射能信号作为中间被控量，与蒸汽被控量一起构成燃烧串级控制系统。这里并不考虑采用预测控制，因为本身所研究的对象有很大的不确定性，加之对象有很强的时变性，预测控制很难达到要求。事实上，对单元机组负荷预测并不重要，而单元机组负荷调节速度才是解决问题的关键。为了便于工程实施和最简单有效的实现系统功能，进行了一些合理的简化和假定。假定锅炉正常运行的燃烧过程近似为具有积分环节、惯性环节、迟延环节等典型环节的线性过程。如果暂时不考虑送风及引风风量等参数的调节，控制系统结构如图 4-1 所示。其中 $S_p(t)$ 为主汽压设定值， $F(t)$ 为燃料量， $P(t)$ 为主汽压力， $\Delta F(t)$ 、 $\Delta P(t)$ 分别为燃料扰动和负荷扰动。

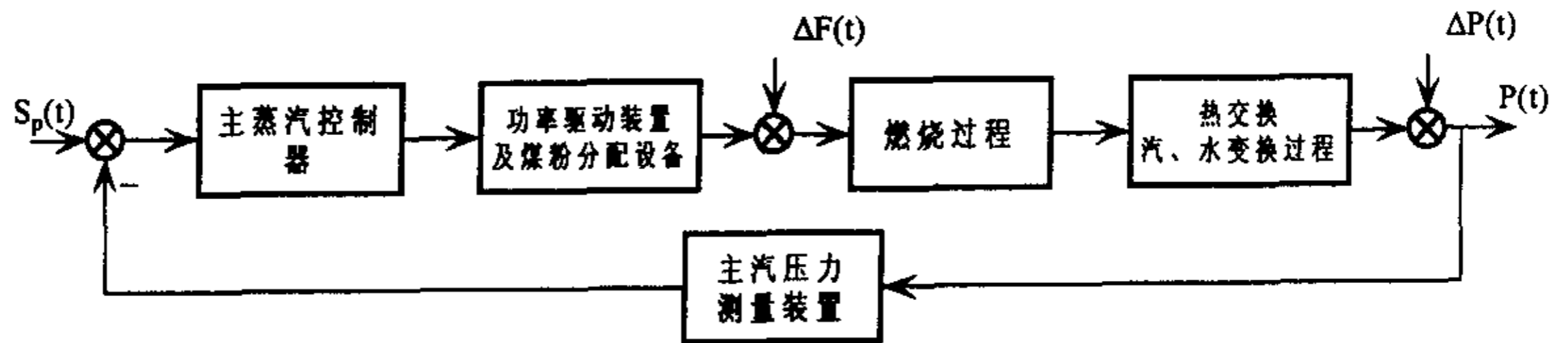


图 4-1 普通主汽控制系统结构图

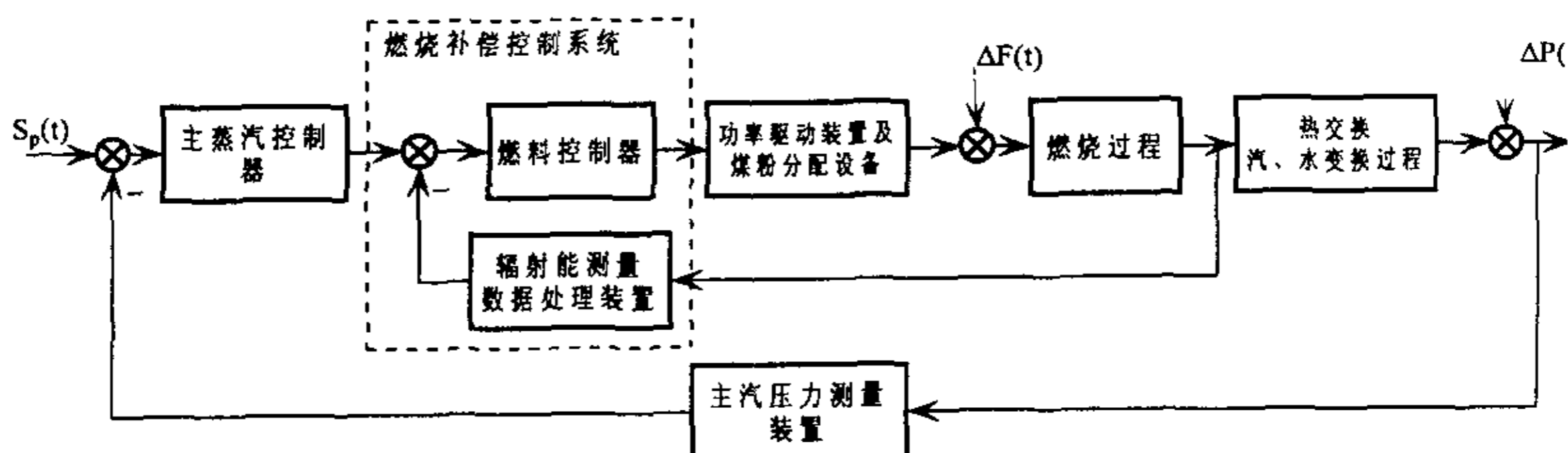


图 4-2 引入燃烧辐射能反馈的主汽控制系统结构图

引入辐射能补偿控制，实际上是在主汽控制回路中加入一个燃料-辐射能控制内部负反馈回路，形成典型的串级控制结构。同样暂时不考虑送风风量等参数的调节以及各控制变量之间的耦合关系，简化结构如图 4-2 所示。由于内部负反馈回路的存在，能有效地克服燃料扰动，及时校正并减小燃料扰动对主蒸汽参数的影响；同时，由于燃烧能量交换和汽水变换过程的惯性时间常数都较大，引入辐射能量反馈回路后，使系统的等效时间常数减小，在系统稳定的前提下主回路增益可以相应提高些，因此可以提高控制质量。辐射能控制内回路的存在还改善了蒸汽压力被调对象的特性。既保证了蒸汽压力具有较快的上升速度，同时几乎不出现明显的超调量，还提高了系统在给定值阶跃扰动下的动态响应能力，加快控制系统跟踪机组负荷变化的速度，改善机组的调峰性能。另外，有效的燃料控制对于燃烧控制经济性的提高和减轻炉管超温现象也是十分有利的，因为它可以最大限度地减小多余的燃料进入炉膛，也避免了炉膛燃烧温度不必要地过高。并且，炉内辐射能信号是炉内燃烧温度水平的直接反映，跟氮氧化物生成密切相关。控制多余的燃料进入炉膛，也就提高了燃料利用率，同时抑制过高的燃烧温度对于控制氮氧化物的生成也是十分有利的。更深入工作将是再结合燃烧温度、过剩空气率及炉内配风方式等运行条件和参数对 NO_x 生成有明显的影响，寻求广泛地也锅炉运行的自动控制系统有机地结合，使锅炉在运行中获得尽可能低的污染物排放水平。

4.2 在双输入双输出模型上的应用

基于上述思想，将火电机组运行过程简化为如下图所示的一个双输入双输出模型：

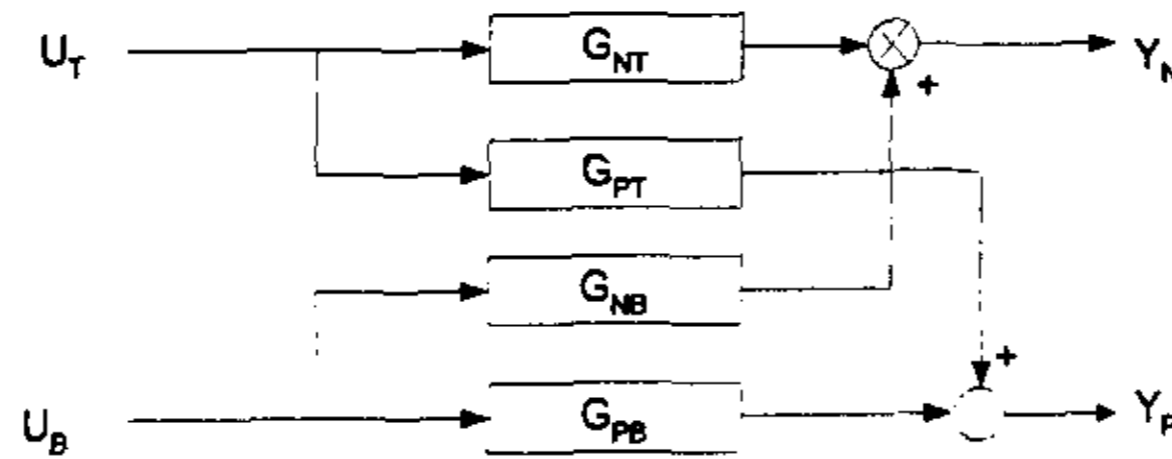


图 4-3 火电机组双输入双输出模型

其中， $G_{NT}(s)$ 为调门开度-机组功率的过程内部传递函数， $G_{PT}(s)$ 为调门开度-蒸汽压力的内部传函， $G_{NB}(s)$ 为给粉指令-机组功率的内部传函， $G_{PB}(s)$ 为给粉指令-蒸汽压力的内部传函。

在下文中，将对模型的结构和阶次做多方面的探讨，其中一共用到五种结构的模型对电厂原始数据进行辨识，并比较辨识结果，最后得到一个最优的模型。这五种模型分别为：

$$\text{模型 1: } G_{NT}(s), G_{PT}(s), G_{NB}(s), G_{PB}(s) = \frac{b_0}{1 + a_0s + a_1s^2} \quad (4.2.1)$$

$$\text{模型 2: } \begin{cases} G_{NT}(s), G_{PT}(s) = \frac{b_0 + b_1s}{1 + a_0s + a_1s^2} \\ G_{NB}(s), G_{PB}(s) = \frac{b_0}{1 + a_0s + a_1s^2} \end{cases} \quad (4.2.2)$$

$$\text{模型 3: } G_{NT}(s), G_{PT}(s), G_{NB}(s), G_{PB}(s) = \frac{b_0 + b_1s}{1 + a_0s + a_1s^2} \quad (4.2.3)$$

Z变换有所不同

$$\text{模型 4: } \begin{cases} G_{NT}(s), G_{PT}(s) = \frac{b_0 + b_1s}{1 + a_0s + a_1s^2} \\ G_{NB}(s), G_{PB}(s) = \frac{k}{1 + Ts} e^{-\tau s} \end{cases} \quad (4.2.4)$$

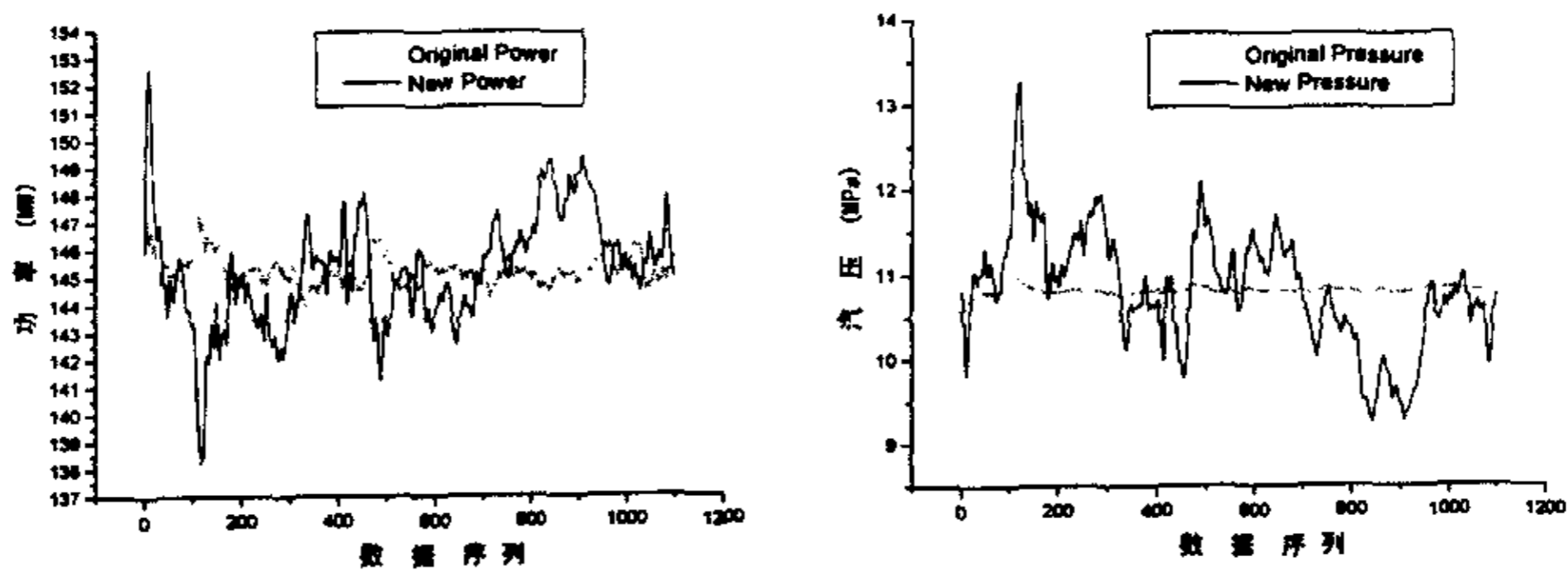
$$\text{模型 5: } G_{NT}(s), G_{PT}(s), G_{NB}(s), G_{PB}(s) = \frac{b_0 + b_1s}{1 + a_0s + a_1s^2} \quad (4.2.5)$$

其中的模型 3 进行 Z 变换的时候, 会把向前回溯的数据个数定为 3 个, 这与其它四个模型利用 (3.2.2) 式进行计算不同, 它们只回溯 2 个数据。

4.2.1 模型 1、模型 2 和模型 3 的辨识结果

在某火电厂 200MW 机组上现场取得九组数据, 每组数据的个数在 400 个到 1000 个不等。

使用模型 1 进行辨识, 结果并不理想, 例如图 4-4 所示的第 1 组数据辨识结果, 其均方差虽然只有 49.5795, 但是根本不能预测输出的趋势:



(a) 机组输出功率

(b) 主汽压力

图 4-4 使用模型 1, 第 1 组数据辨识结果预测输出与实际输出比较, 均方差之和为 49.5795

而使用模型 2 对第 7 组数据辨识的结果更差, 均方差达到 878.238, 这是根本不能容忍的结果。其预测输出拟合图如下图所示。所以这个模型结构也是不适用于火电机组双输入双输出过程的系统辨识的。

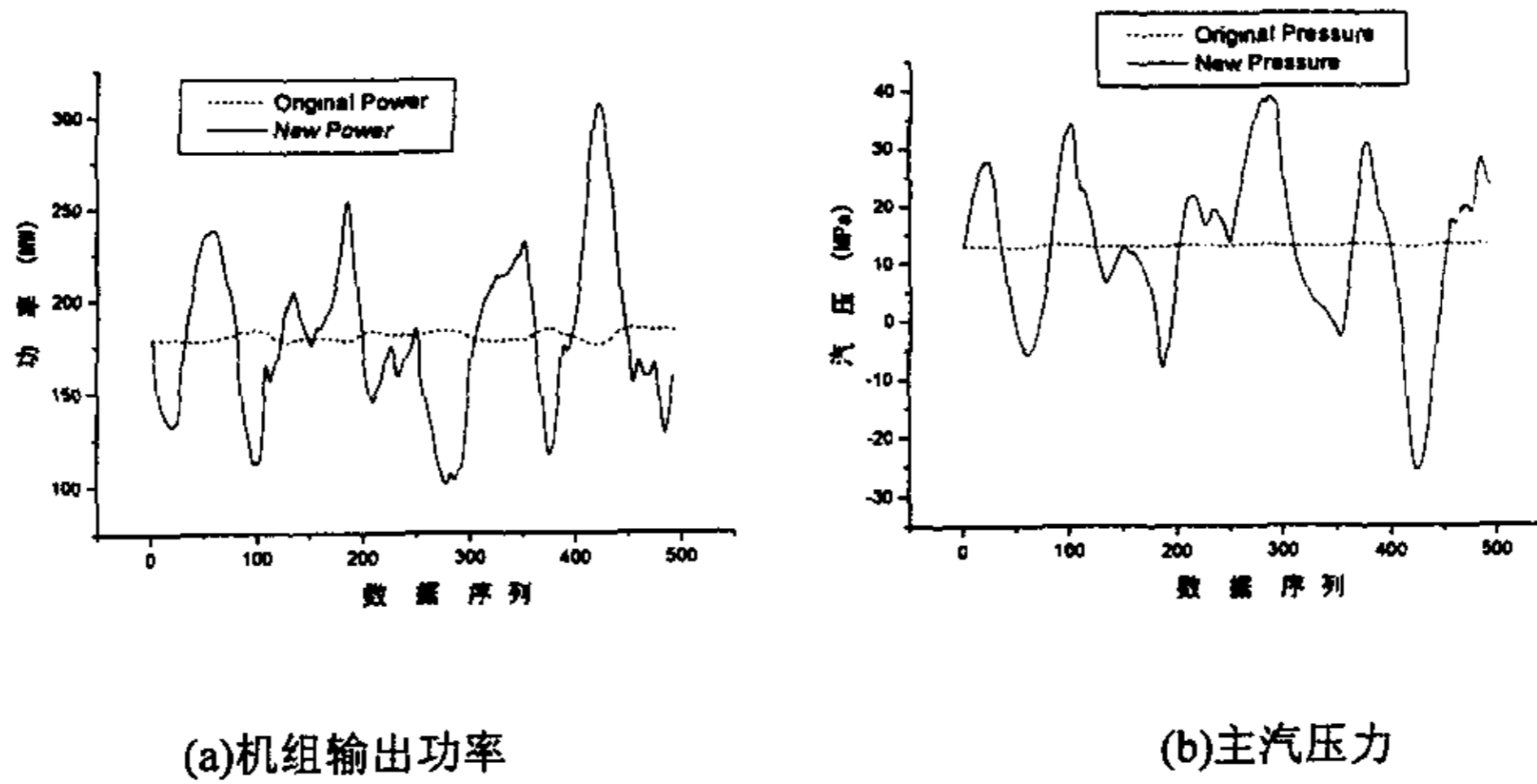


图 4-5 使用模型 2，第 7 组数据辨识结果预测输出与实际输出比较，均方差之和为 878.238

再用模型 3，对第 9 组数据进行辨识，情况稍微好一些，均方差为 68.2845；但是其预测输出拟合图说明这也是不能作为可用的结果的。如图 4-6 所示。

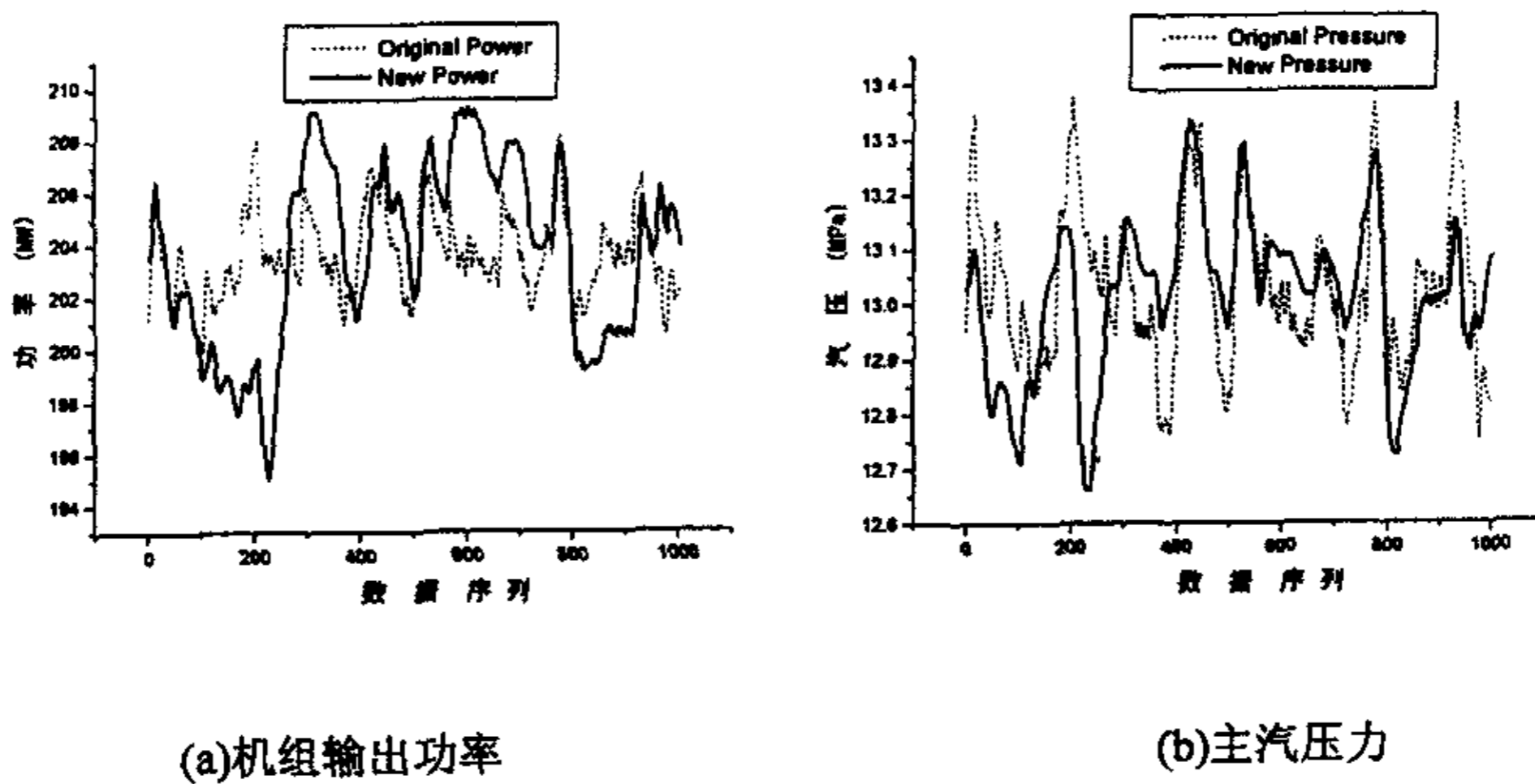


图 4-6 使用模型 3，第 9 组数据辨识结果预测输出与实际输出比较，均方差之和为 68.2845

综合以上情况，可以认为模型 1、2、3 都是不适合用作这样一个双输入双输出系统的。下面本文讨论模型 4 的辨识情况。

4.2.2 模型 4 的辨识结果

本文之所以要使用模型 4 的结构, 是要试图使模型结构能够直接反映过程的动态特性。同时, 由于调门开度的变化几乎是可以立即反应到机组功率和汽压的变化的, 从调门开度到功率和汽压的这两个内部传函仍旧沿用模型 2 的结构, 而从给粉指令到功率和汽压则是两个带有纯延迟大滞后性质的过程, 因而将带滞后的一阶模型结构应用到从给粉指令到功率和到汽压这两个内部传函上。在 (4.2.4) 式中, k 为比例系数, T 为惯性时间, τ 为延迟时间。其中 k 可以直观的反映给粉和辐射能的静态比例, τ 为延迟时间。

首先要根据原始数据把响应延迟时间确定下来。本文在电厂现场数据中, 初步确定 τ 为 36 个采样时间, 即为 180 秒。然后再对某一组数据进行辨识试验, 变且不断变换 τ 的值, 直到得到一个最优的 τ 值使得均方差之和最小。本文对第七组数据的试验计算过程如表 4-1 所示。

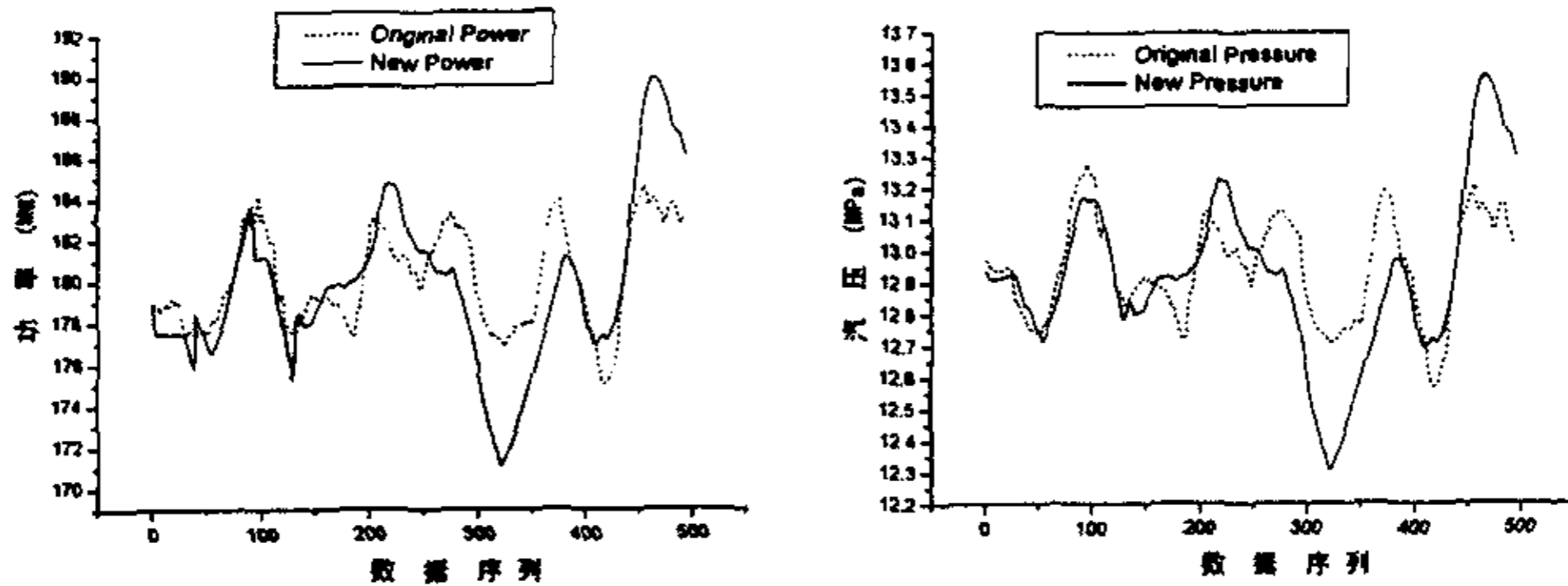
表 4-1 确定最优 τ 值的计算过程

计算步数	τ 值与采样时间的比值	辨识均方差之和
1	36	58.2762
2	35	57.7293
3	33	56.8099
4	30	55.8957
5	25	56.5278
6	27	55.9795
7	28	55.7916
8	29	55.8519

当计算进行到第四步, 发现 τ 的与采样时间的比值减小到 30 时, 均方差仍在不断减小, 于是再从 25 开始, 发现均方差比较大, 再从 25 开始增大, 最后得出 τ 与采样时间的比值在 28 左右时有最优的辨识结果。后面对另一组数据的辨识也将采用这个值。

但是辨识结果并不能让人满意, 如对第 7 组数据的辨识均方差达到 58.2503, 而对第 9 组数据的辨识均方差更是达到了 109.287, 它们的辨识

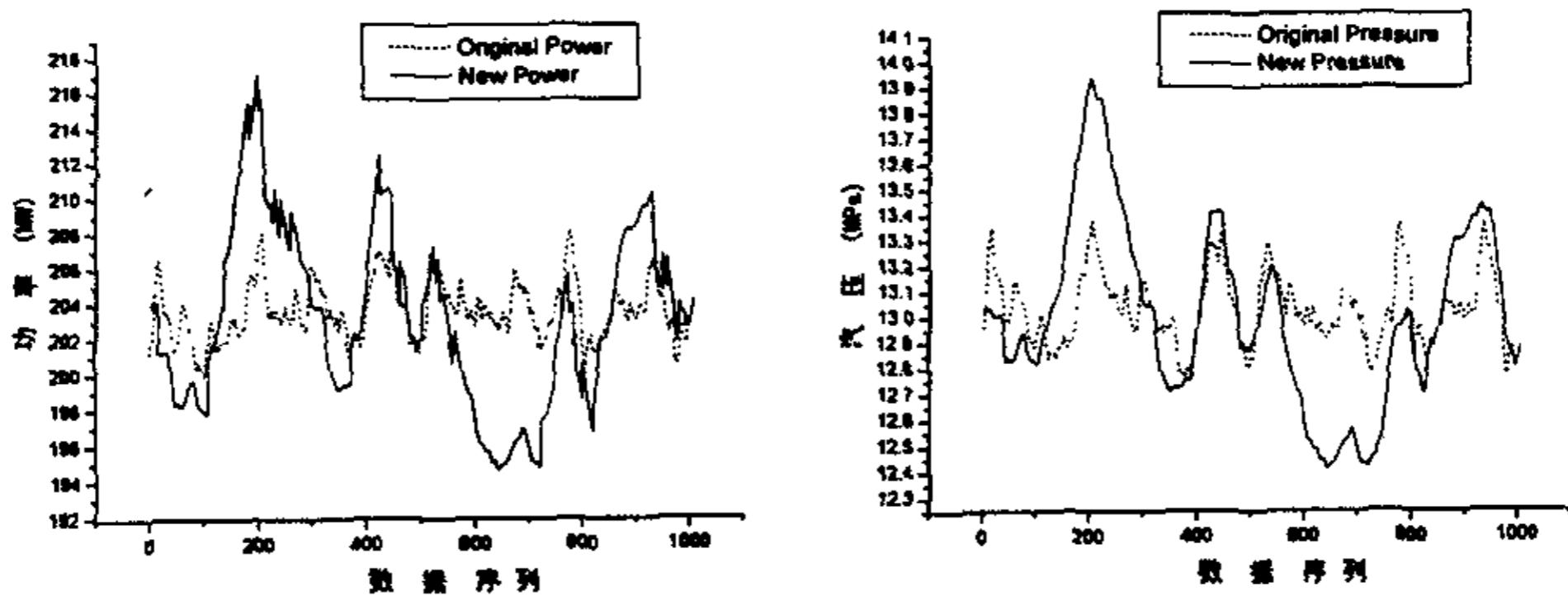
结果预测输出拟合图如下:



(a) 机组输出功率

(b) 主汽压力

图 4-7 第 7 组数据辨识结果预测输出与实际输出比较, 均方差之和为 55.7916



(a) 机组输出功率

(b) 主汽压力

图 4-8 第 9 组数据辨识结果预测输出与实际输出比较, 均方差之和为 107.558

这一模型结构不适用的原因, 与将火电机组简化为双输入双输出过程有关, 因为在实际运行中还有很多影响生产过程的因素没有被考虑进来, 从给粉到机组功率和蒸汽压力是一个复杂的过程, 其中具有一定的非线性因素, 同时还有很多未知的影响因素, 甚至在有的工况下这些本文没有考虑的因素占据了主导地位, 这些都是导致模型 4 失败的原因。

4.2.3 模型 5 的辨识结果

本文使用模型 5, 得到了令人满意的结果。下面将会详细叙述计算结果。如前所述, 已经在某电厂 200MW 机组上取得了九组现场数据。使用

华中科技大学硕士学位论文

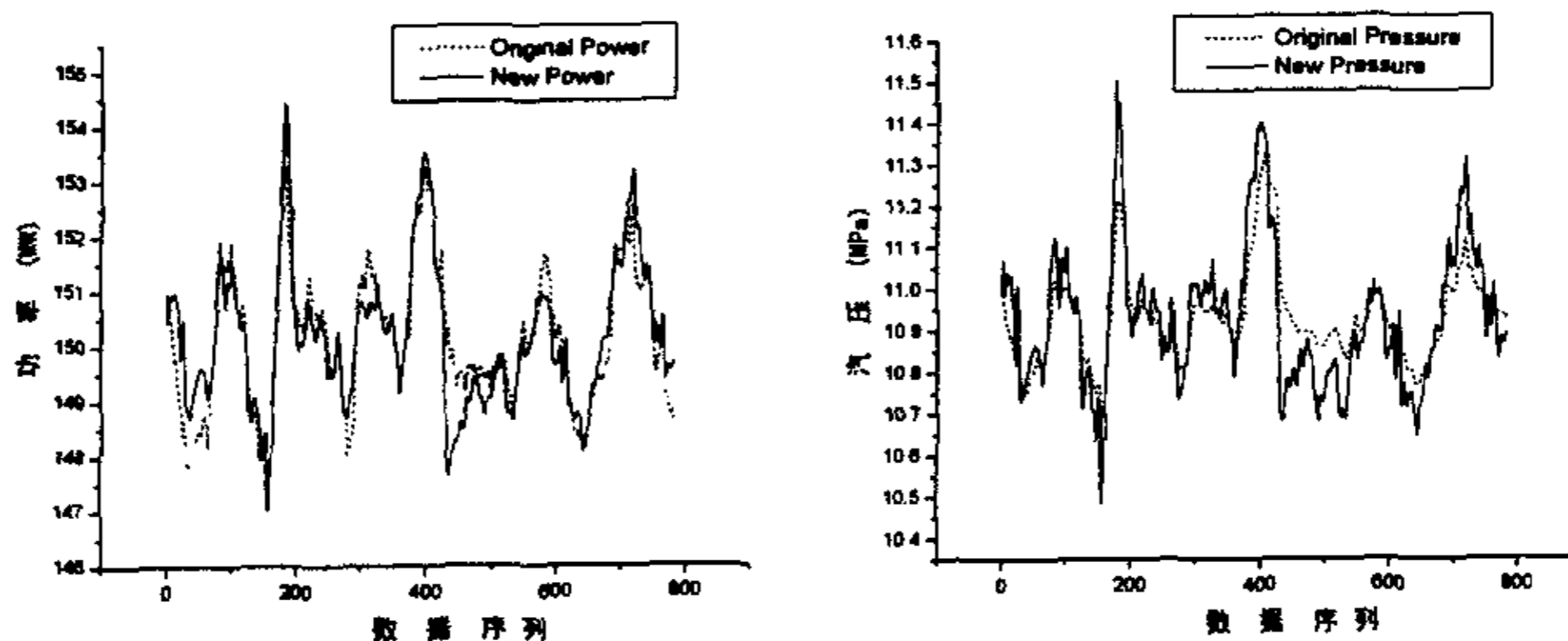
模型 5, 得到辨识结果如表 4-2 所示。

表 4-2 使用模型 5, 九组双输入双输出数据辨识结果

	$G_{NT}(s)$ b ₀ 、b ₁ 、a ₀ 、 a ₁	$G_{PT}(s)$ b ₀ 、b ₁ 、a ₀ 、 a ₁	$G_{NB}(s)$ b ₀ 、b ₁ 、a ₀ 、 a ₁	$G_{PB}(s)$ b ₀ 、b ₁ 、a ₀ 、 a ₁	均方差 之和
Data1	24.2217 123554 64862 132035	-0.209473 -13.477 153.351 0.284401	0.010078 -33.9132 412.997 11389.3	0.0015039 -2.57139 420.953 5732.34	5.9001
Data2	4.51521 1270.03 642.333 3755.87	-0.0613542 -1.09458 68.5365 -1.85526	-0.0521991 -20.2983 247.015 4516.5	0.00521029 -3.64185 370.182 2216.78	56.9764
Data3	3.23748 -140.15 820.126 -176.782	0.285638 -203.194 736.507 6538.97	0.00210958 -7.87096 170.339 2660.8	0.0012378 -1.54472 308.118 2537.35	14.3965
Data4	0.966468 -32.3804 404.286 -211.009	-0.228264 -130.274 911.84 45.5366	-0.0213056 -18.2887 212.131 4863.72	0.00269237 -2.59338 352.028 3509.51	26.5689
Data5	-2.85093 1.77799 497.173 -288.621	-0.170409 0.306538 21.0486 3.4245	0.0431771 -22.0221 314.142 5206.08	0.00404813 -2.80168 322.662 2620.64	12.2508
Data6	4.22878 2.398e+006 301091 972460	-0.0696526 86.3904 299.68 0.0408652	-0.0277671 -13.3642 143.788 3569.27	0.00654872 -3.76434 305.57 1790.41	30.4281
Data7	4.43802 1489.84 493.396 2010.98	-0.0285182 14.2581 345.217 9157.4	0.0272759 -16.7055 210.927 8297.9	0.00262955 -1.68092 283.888 5406.38	4.6642
Data8	1.28272 -99.7443 499.72 10870.5	-0.0079934 -12.4242 154.701 23.8817	0.0017852 -13.949 168.256 5775.77	-0.0012348 -0.99513 187.112 5538.87	4.6970
Data9	2.41244 117.189 131.505	-0.0093923 -4.95587 124.065	-0.0021238 -12.839 225.146	0.00046278 -1.03694 210.292	8.3259

	813.42	1007.13	4115.67	4048.24
--	--------	---------	---------	---------

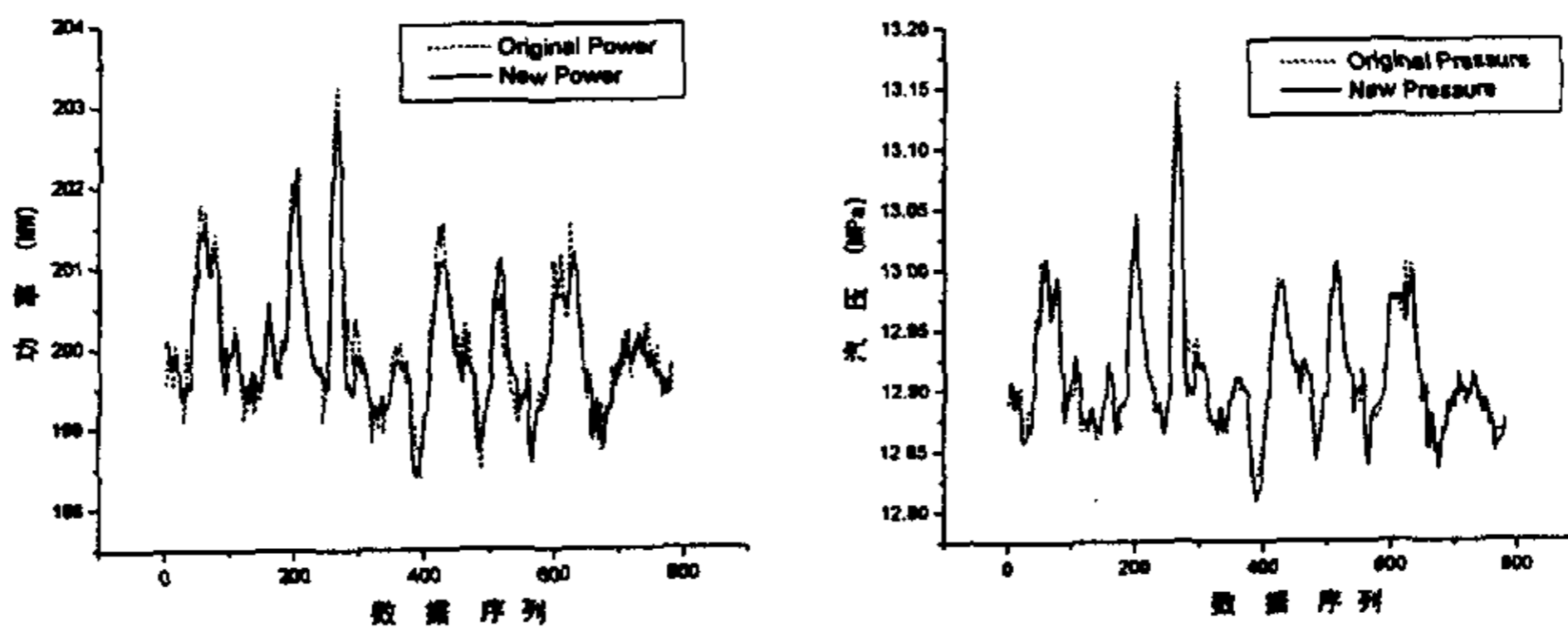
任取其中的两组模型，如第 5 组和第 8 组数据辨识得到的模型，其辨识结束时的预测输出与实际输出比较图如 4-9 和 4-10 所示。可以看出，尽管从均方差看起来二者相差很大，可是在曲线的拟合上并不能看出明显差别。



(a) 机组输出功率

(b) 主汽压力

图 4-9 第 5 组数据辨识结果预测输出与实际输出比较，均方差之和为 12.2508



(a) 机组输出功率

(b) 主汽压力

图 4-10 第 8 组数据辨识结果预测输出与实际输出比较，均方差之和为 4.6970

同样，仅仅得到这些结果还远远不够，还必须对 9 个模型进行交叉校验。在校验的过程中，发现第 2、3、4、5 组数据得到的模型对其它数据进行校验时出现了奇异值，所以要将其抛弃。将剩下的五组模型求算术平均值，得到如下式所示的模型：

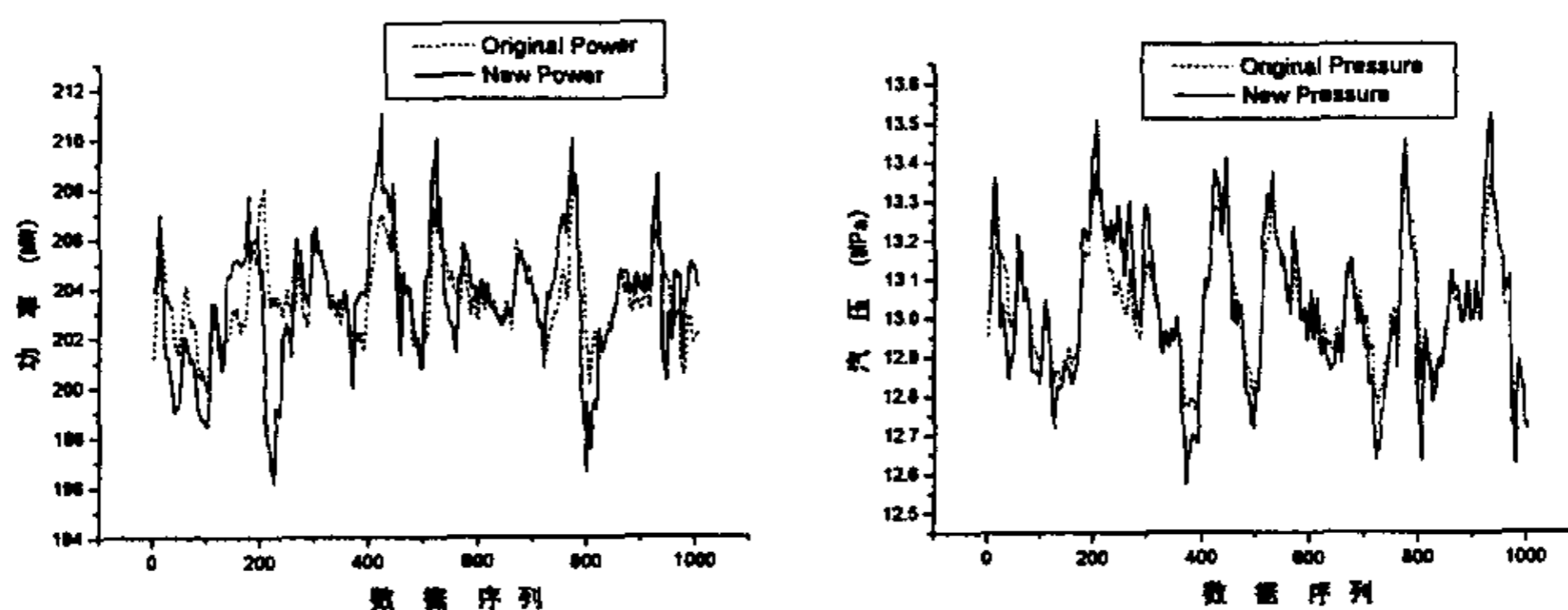
$$\begin{cases} G_{NT}(s) = \frac{7.31673 + 504642s}{1 + 73415.5s + 223638s^2} \\ G_{PT}(s) = \frac{-0.0650059 + 13.9583s}{1 + 215.403s + 2037.75s^2} \\ G_{NB}(s) = \frac{0.00184978 - 18.1542s}{1 + 232.223s + 6629.58s^2} \\ G_{PB}(s) = \frac{0.00198202 - 2.00974s}{1 + 281.563s + 4503.25s^2} \end{cases} \quad (4.2.6)$$

用这个模型对全部九组数据进行校验，得到的预测输出与实际输出的均方差如表 4-3 所示：

表 4-3 统一模型校验结果

	数据个数	辨识均方差之和	校验均方差之和
Data1	1100	5.9001	6.27067
Data2	1000	56.9764	129.677
Data3	836	14.3965	50.3972
Data4	1068	26.5689	179.999
Data5	782	12.2508	35.7231
Data6	894	30.4281	29.1981
Data7	493	4.6642	7.27212
Data8	782	4.6970	34.3049
Data9	1004	8.3259	36.5934

可以看出，要想得到一个统一的对所有数据都适用的模型，是要付出代价的，这个代价就是拟合精度。不过，单纯的看均方差是不能说明问题的，例如，画出统一模型对第九组数据的校验输出与实际输出的比较图如图 4-11 所示，虽然其拟合均方差也达到了 36.5934，但是对于趋势的预测却很精确：



(a) 机组输出功率

(b) 主汽压力

图 4-11 第 9 组数据校验预测输出与实际输出比较，均方差之和为 36.5934

4.3 结果分析与本章小结

正如本文第二章所描述的那样，当对若干组数据进行辨识得到结果之后，工作还远远没有完成，还必须用每一组数据得到的模型对其它组数据进行交叉校验。在这个过程中需要抛弃那些产生了奇异值的模型。这一步工作完成后，再对剩下适用的模型对应参数求算术平均值，以使得到的统一模型具有普适性，即对不同工况下的数据具有大致相同的精确度。这一点是很重要的，否则模型只对某些工况适用的很好，而对另一些工况下的数据则不适用，或者误差很大，这是我们所不允许的。最后，还要用这个统一模型再对所有数据进行校验，如果全部的计算表明，这个模型是合适的，这时候辨识工作才算完成。

至此，本文已经将本章第一节所描述的辐射能反馈信号的火电机组负荷控制系统中的双输入数学模型辨识出来了，它对现场数据的预测输出拟合得非常好，可以认为这些模型是能够反映过程的本身特性的。

5 在火电机组单输入单输出过程中的应用

本文在上一章介绍了采用辐射能反馈信号的火电机组负荷控制系统，并基于这种思想，把多输入多输出的系统辨识方法应用到给粉、调门开度到机组功率和蒸汽压力这样一个双输入双输出系统中，文章最后得到了一个很精确的模型，经过交叉检验和阶跃响应试验，都说明这是一种行之有效的系统辨识方法。在这一章中，将继续把这种系统辨识方法应用到火电机组负荷控制系统中去，辨识从给粉指令到辐射能信号这样一个单输入单输出过程的数学模型。

利用在某电厂连续采集的给粉指令和辐射能信号来作为过程输入数据，其中辐射能信号是华中科技大学创能科技有限公司开发的 ZHC-A 型燃煤锅炉火焰图像监视系统所产生的反映炉膛内燃烧状况的参数，采用辐射能反馈信号的火电机组负荷控制系统在前文已经作过介绍。

5.1 采用二阶模型的给粉-辐射能过程辨识

一共有六段数据，分别为机组不同负荷下的数据。采用的模型传递函数的结构为：

$$G(s) = \frac{B(s)}{E(s)} = \frac{b_0 + b_1s}{1 + a_0s + a_1s^2} \quad (5.1.1)$$

式中 $B(s)$ 为给粉指令函数， $E(s)$ 为辐射能函数。这六段数据的分别辨识结果如表 5-1 所示：

华中科技大学硕士学位论文

表 5-1 采用二阶模型的六组数据辨识结果

	负荷 (MW)	b_0	b_1	a_0	a_1	均方差
Data1	210	0.0106917	47.6499	372.161	67902.4	104.593
Data2	110	0.242399	459.661	1837.94	39361.8	39.616
Data3	210	0.493367	111.588	1651.17	141397	66.2292
Data4	170	0.0220379	-7.51998	8.89427	6307.76	151.96
Data5	170	0.0215008	-7.02592	31.8676	-0.207651	56.9506
Data6	160	0.122092	-8.22632	36.3368	-0.573125	44.0105
average	—	0.152015	99.3544	656.395	42494.7	—

在辨识程序的迭代计算完成后, 把依据辨识结果得到的预测输出与原始数据中的辐射能数据进行比较, 如图 5-1、5-2、5-3、5-4、5-5、5-6 所示。

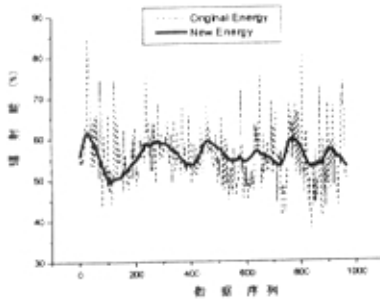


图 5-1

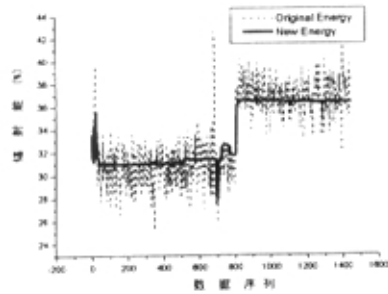


图 5-2

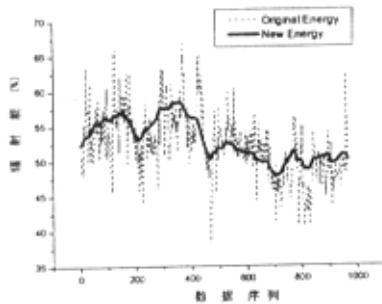


图 5-3

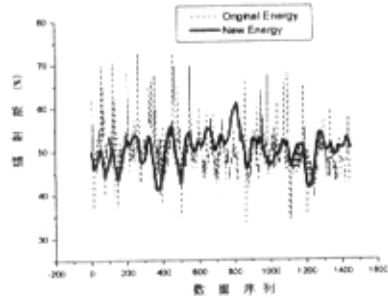


图 5-4

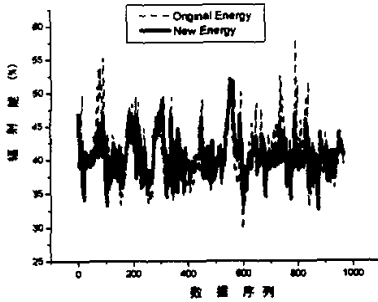


图 5-5

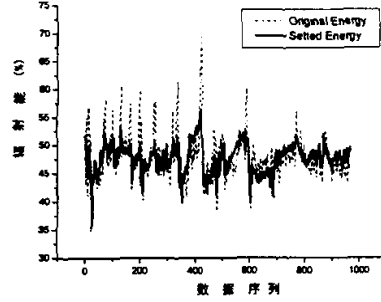


图 5-6

图 5-1~图 5-6 六组数据辨识结果

可以看出根据辨识结果所得到的预测输出与原始数据吻合的很好。并且，对各个相应的参数取算术平均值，如表 5-1 最后一行所示，这样可以得到一个应该对所有数据都适用的模型如下式所示：

$$G(s) = \frac{B(s)}{E(s)} = \frac{0.152015 + 99.3544s}{1 + 656.395s + 42494.7s^2} \quad (5.1.2)$$

用这个模型对所有六组数据进行校验，得到的预测输出与实际数据的均方差可见表 5-2 所示。

表 5-2 辨识均方差及两个方法校验均方差比较

	数据个数	辨识均方差	式 (5.1.2) 模型校验均方差	式 (5.1.3) 模型校验均方差
Data1	965	104.593	160.364	161.225
Data2	1447	39.616	56.2938	51.7735
Data3	965	66.2292	84.9883	82.5444
Data4	1447	151.96	157.36	161.467
Data5	965	56.9506	139.167	144.846
Data6	965	44.0105	57.4666	57.4826

但是，注意到各组模型参数相差比较大，就有必要进行交叉检验，即使用第一组的模型来对其它五组的数据进行校验。这样可以剔除个别虽然可以比较好地吻合原始输出数据，但是本身不具有物理意义的模型。一共

要进行 $6 \times 5 = 30$ 次校验，结果发现，从第一组到第四组的模型在对其它组数据进行校验时都可以得到令人满意的结果，但是第五组和第六组模型就出现了奇异值。下面图 5-7 和 5-8 分别是第五组模型对第四组数据校验结果图和第六组模型对第四组数据校验的结果图，可以看到在校验的后段，预测数据呈现极度发散的状态，这说明模型是错误的。我们认为，只要模型对任何一组数据的校验不符合要求，就必须抛弃这个错误的模型。实际上，观察表 5-1 中各个模型的对应参数就可以看出，模型五和模型六是明显错误的两个模型。

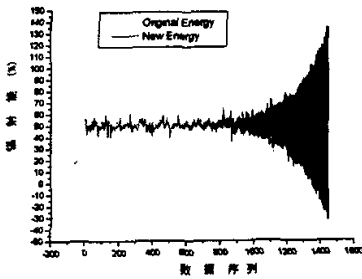


图 5-7 模型五对数据四的校验，
均方差为 833.512

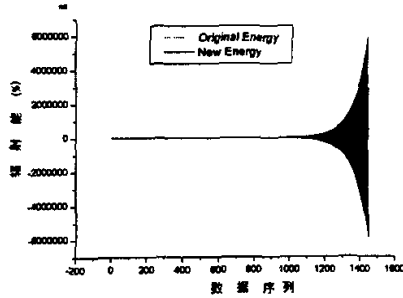


图 5-8 模型六对数据四的校验，
均方差为 $3.7915e+007$

于是取模型一到模型四的各参数的算术平均值，得到如下式所示的模型：

$$G(s) = \frac{B(s)}{E(s)} = \frac{0.192124 + 152.845s}{1 + 967.541s + 63742.2s^2} \quad (5.1.3)$$

用这个模型对六组数据进行校验，所得到的均方差见表 5-2 所示，其拟合结果如图 5-9、5-10、5-11、5-12、5-13、5-14 所示。

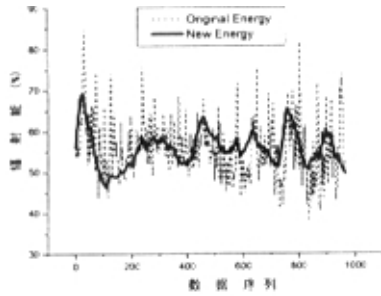


图 5-9

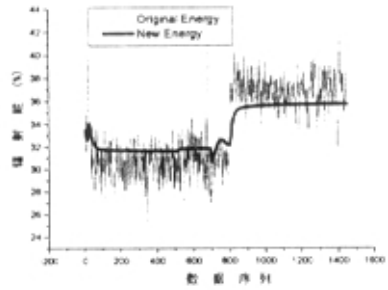


图 5-10

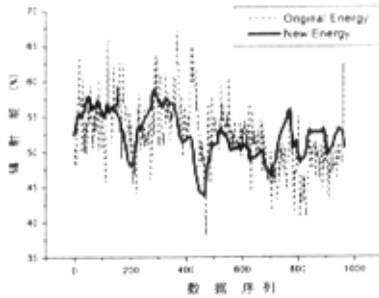


图 5-11

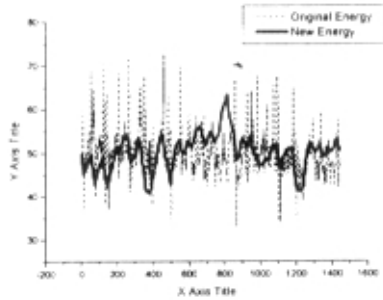


图 5-12

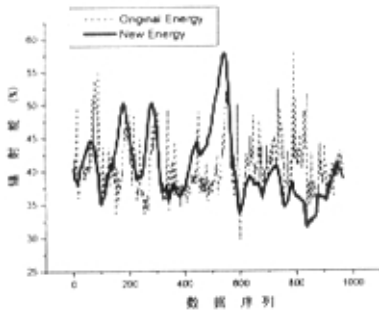


图 5-13

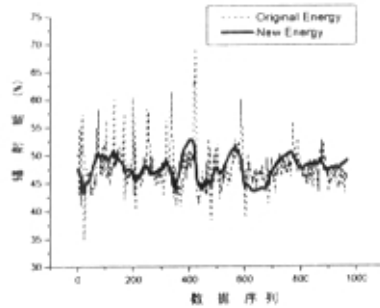


图 5-14

图 5-9~图 5-14 二阶模型六组数据校验结果

可以看出，用上面所述方法进行处理后得到的辨识结果对六组数据进行校验，其均方差与拟合效果图都是令人满意的。同时，还抛弃了错误的模型，避免出现奇异值的现象。这些结果都证明了本文提出的辨识方法是正确高

效的。

5.2 采用一阶模型的给粉-辐射能过程辨识

在上文中已经得到了一个比较精确的模型，并且也通过了校验检查。但是，由于其模型结构并不能简单地看出过程的物理意义，我们就尝试使用如下的模型结构形式来对给粉-辐射能过程进行辨识：

$$G(s) = \frac{B(s)}{E(s)} = \frac{k}{1+Ts} e^{-\tau s} \quad (5.2.1)$$

式中， k 为比例系数， T 为惯性时间， τ 为延迟时间。其中 k 可以直观的反映给粉和辐射能的静态比例， τ 为响应延迟时间。

首先需要根据原始数据把响应延迟时间确定下来。任取一段原始数据如图 5-2-1 所示。由于在辨识过程中需要对采样时间做除法，所以这里的 τ 可以表示为采样时间的整数倍。反映到图 5-15 中，就是给粉指令中的 a 点和辐射能中的 b 点两个峰值所相距的数据个数，这个数据个数就是 τ 值与采样时间的比。通过图 5-15，初步确定这个比例为 18，再对第一组数据进行辨识试验，最后确定使辨识均方差最小的 τ 值即为 18。计算过程见表 5-3。

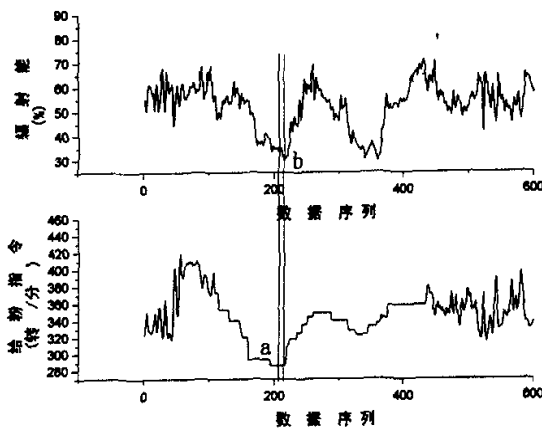


图 5-15 确定 τ 与采样时间的比例值

表 5-3 确定最优 τ 值的计算过程

计算步数	τ 值与采样时间的比值	辨识均方差之和
1	18	150.886
2	19	151.192
3	25	156.647
4	17	151.132
5	10	153.106

从上表的计算过程中可以知道, τ 取为采样时间的 18 倍时有最优解。同样使用上文使用的六组数据, 得到它们的辨识计算结果如表 5-4 所示。

表 5-4 改变模型结构后六组数据辨识结果

	机组负荷 (MW)	b ₀	b ₁	均方差
Data1	210	0.104557	11.6359	150.886
Data2	110	0.28217	29.2188	44.8095
Data3	210	0.120097	61.483	86.2315
Data4	170	0.0483365	14.6285	136.101
Data5	170	0.0487154	14.8394	86.0507
Data6	160	0.114934	27.4242	53.3747
average	—	0.119802	26.5383	—

限于篇幅, 这里就不再给出各组数据的辨识结果拟合图。

当分别用这六个模型对其它五组数据进行交叉校验的时候, 我们发现没有如上文出现的情况那样存在错误的模型, 也就是说这六个模型分别对其它五组模型的交叉校验都是适用的。

同样, 我们采用求算术平均值的办法, 得到下式所示的模型:

$$G(s) = \frac{B(s)}{E(s)} = \frac{0.119802}{1 + 26.5383s} e^{-90s} \quad (5.2.2)$$

用这个模型对六组数据进行校验, 得到的均方差如表 5-2-3 所示。

表 5-5 两种不同结构的模型辨识精度比较

	数据个数	二阶模型辨识均方差	二阶统一模型校验均方差	一阶模型辨识均方差	一阶统一模型校验均方差
Data1	965	104.593	161.225	150.886	151.443
Data2	1447	39.616	51.7735	44.8095	65.861
Data3	965	66.2292	82.5444	86.2315	86.5367
Data4	1447	151.96	161.467	136.101	147.278
Data5	965	56.9506	144.846	86.0507	112.941
Data6	965	44.0105	57.4826	53.3747	53.745

图 5-15、5-17、5-18、5-19、5-20、5-21 为 (5.2.2) 式模型对六组数据校验的预测输出与实际输出比较图。

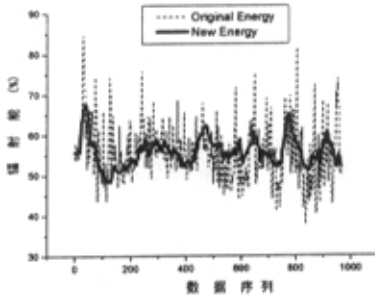


图 5-15

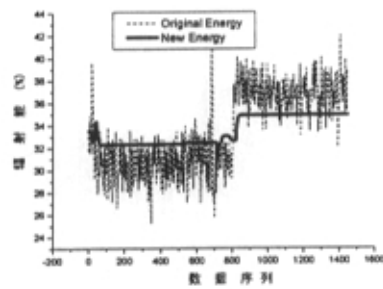


图 5-16

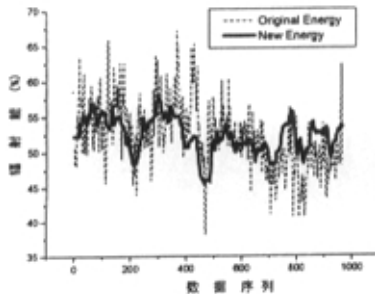


图 5-17

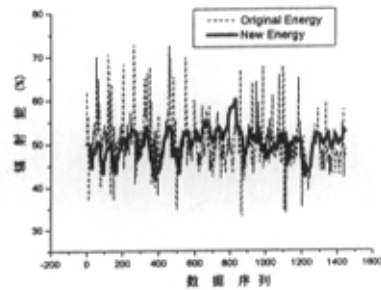


图 5-18

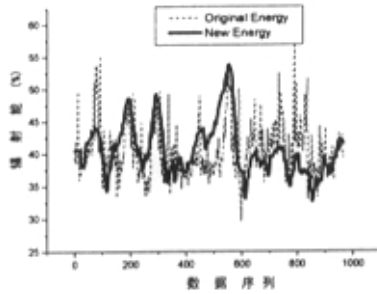


图 5-19

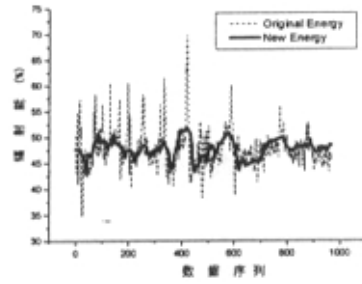


图 5-20

图 5-15~图 5-20 一阶模型六组数据校验结果

5.3 结果分析

下面将分别对本章用到的两个给粉-辐射能模型进行比较分析,以及把前一章得到的双输入双输出模型与本章得到的一阶带滞后的给粉-辐射能模型进行比较,得到有意义的结果。

5.3.1 给粉-辐射能过程辨识中两个模型比较分析

在本章的计算中,一共用到两种结构的模型,一是如(5.1.1)式所示的二阶模型,二是如(5.2.1)式所示的带滞后的一阶模型。表 5-2-3 中列出了采用 2 阶模型进行辨识得到的均方差、得到统一模型进行校验得到的均方差、采用 1 阶模型进行辨识得到的均方差和得到统一模型后进行校验得到的均方差。从这个表中可以看出两种不同结构的模型,其辨识精度也是不一样的。

虽然在辨识结束时,得到的预测输出与实际输出的均方差这一项上,采用(5.2.1)式所示的有滞后的一阶模型进行辨识时,在有些数据的计算结果上要比采用如式(5.1.1)所得到的均方差要大一些,可是采用(5.2.1)式的模型进行给粉-辐射能过程辨识时,不会产生错误的模型。采用(5.2.1)式模型的另一个优点在于,最终得到的模型是将几组数据辨识得到的模型

相应参数取算术平均值后得到的，这个模型才是最后的结果。从上表中可以看出，采用(5.2.1)式模型结构得到的最终统一模型对六组数据的校验结果明显要比采用二阶模型结构的要好。并且，对两个模型作阶跃响应试验如图5-21所示，可以看出，采用加滞后的一阶模型要比二阶模型响应时间更短。

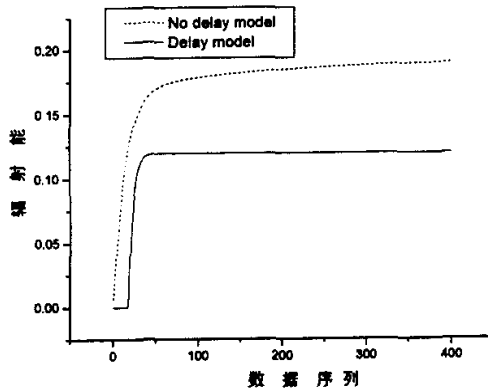


图 5-21 给粉-辐射能过程，两种结构模型阶跃响应对比

通过以上分析，可以将(5.2.2)式的模型作为最终的给粉-辐射能过程的数学模型。

至此，本文已经把电厂火电机组负荷控制系统中的一个至关重要的过程的数学模型辨识出来。有了给粉-辐射能过程的数学模型，就可以将给粉-辐射能这一过程加入原有的负荷控制系统中，作为一个负反馈回路。有了这个回路，就可以快速地反映炉膛里燃烧状态的变化，从而可以快速调节，实现优化控制。

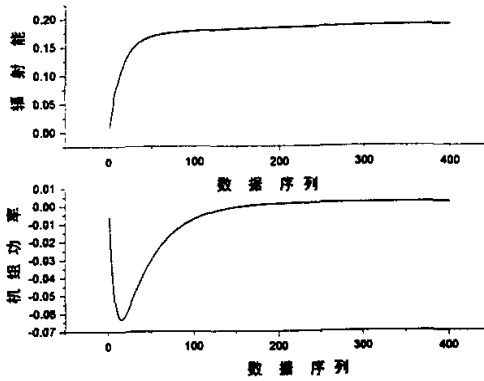
5.3.2 双输入双输出模型与单输入单输出模型的分析

在第四章第一节中，本文描述了采用辐射能反馈信号的火电机组负荷控制系统，其基本思想就是辐射能信号可以比蒸汽压力和机组功率更迅速

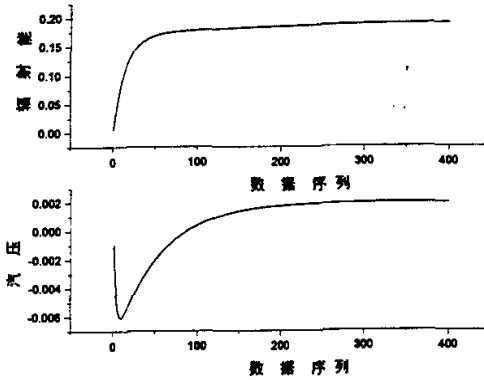
华中科技大学硕士学位论文

更准确地反映出炉膛里燃烧状况的变化。这里将把上一章得到的给粉指令、调门开度到机组功率和蒸汽压力这样的双输入双输出过程的数学模型与本章计算得到的给粉-辐射能的单输入单输出过程模型进行比较,从辨识结果的角度说明这种控制思想提出的正确性。

将第四章(4.2.6)式所示的模型中给粉到功率和给粉到汽压这两个内部传函进行阶跃响应试验,并将之与(5.1.3)式的给粉到辐射能的带滞后一阶模型进行比较,得到结果图如下:



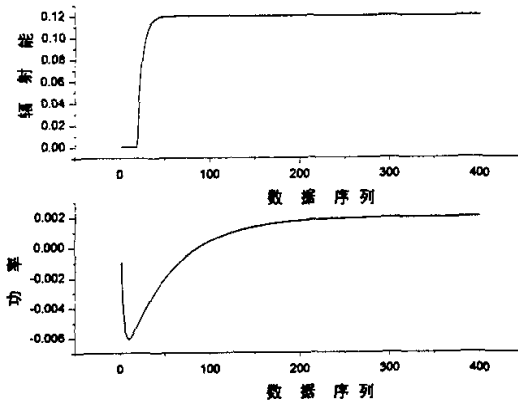
(a) 给粉-辐射能与给粉-机组功率



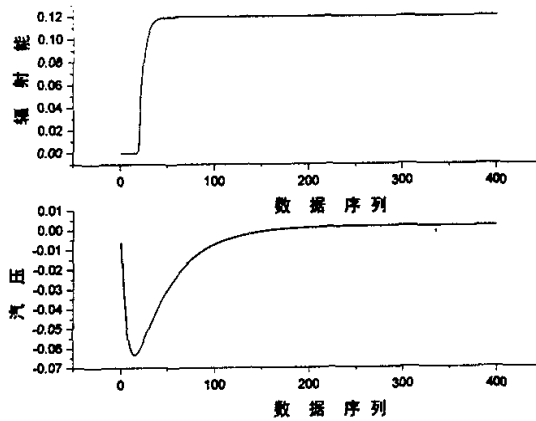
(b) 给粉-辐射能与给粉-蒸汽压力

图 5-22 二阶单输入单输出模型与双输入双输出模型阶跃响应曲线比较

从这里,可以很明显地看出本文辨识得到的给粉-辐射能的模型要比给粉-机组功率和给粉-蒸汽压力这两个模型的响应时间要短得多,这还是并不理想的给粉-辐射能的模型得到的结果。使用(5.2.2)式的带滞后的一阶给粉-辐射能模型作这样的比较,如下图所示。



(a)给粉-辐射能与给粉-机组功率



(b) 给粉-辐射能与给粉-蒸汽压力

图 5-23 一阶单输入单输出模型与双输入双输出模型阶跃响应曲线比较

这里辐射能反映炉膛燃烧状况变化的优势更加突出了,这就给本文第四章第一节所描述的采用辐射能反馈信号的火电机组负荷控制系统的思想提供了有力的佐证。采用辐射能反馈信号的火电机组负荷控制系统,由

于引入了给粉指令反应迅速的辐射能信号作为中间被调量，可以极大改善控制品质，提高控制响应速度，从而简单而高效地实现电站锅炉燃烧控制。

5.4 本章小结

本章继续讨论第四章提出的系统辨识方法在采用辐射能反馈信号的火电机组负荷控制系统中的应用。到这里，本文已经将其中所有过程的数学模型辨识出来了，这些模型都经过了自身辨识、交叉校验、求算术平均值、统一模型再校验这样几个步骤，对现场数据的预测输出拟合得非常好。

同时，本文还将在计算给粉-辐射能的数学模型时所试验的两个模型进行了阶跃响应试验，由此知道如式(5.2.1)所描述的带滞后的一阶模型结构对于火电机组中从给粉到辐射能这一环节更加适用。最后还将两个单输入单输出模型与第四章得到的双输入双输出模型中与给粉指令相关的两个内部传函进行了对比，发现无论给粉-辐射能的过程采用什么样的模型结构，其延迟时间都要比给粉指令-机组功率和给粉指令-蒸汽压力这两个过程的延迟时间要短得多。这就从模型辨识的角度说明了采用辐射能反馈信号的火电机组负荷控制系统思想的正确性。

6 全文总结及未来工作展望

目前的系统辨识方法一般需要在系统输入端加入激励信号，而且多输入多输出系统的在线辨识仍很困难。本文提出一种基于牛顿迭代法的多输入、多输出对象模型迭代辨识方法，模型参数更新的依据是使模型预测输出与全部采样时刻的对象实际输出之间的均方差递减，直到收敛。这种基于全局数据迭代的辨识方法可进行闭环辨识，无需外加激励信号，适用于多输入多输出对象的在线辨识。

本章将就全文进行总结，同时对下一步的工作提出建议和构想。

6.1 全文总结

(1) 本文首先系统的总结了目前国内外在系统辨识领域内的主要方法和理论，然后在此基础上，比较全面的总结了前人的主要研究成果和结论。在介绍了系统辨识的基本概念和基本原理之后，本文重点介绍了最小二乘类的系统辨识方法。其中包括最小二乘问题的解，即有一次完成算法和递推算法之分别。在当今计算机被广泛利用的前提下，获得主要应用的是递推算法。另外还有依模型阶次的递推算法，主要用于确定模型的最佳阶次。随后还介绍了最小二乘类的其它几种算法，如适应算法、偏差补偿最小二乘法、增广最小二乘法、广义最小二乘法等等它们适用于不同情况下的系统辨识问题。

(2) 本文在广泛阅读国内外相关文献以后，提出一种新的基于牛顿迭代法的全局数据迭代的系统辨识方法。文章在第三章重点介绍了这种辨识方法。全局数据迭代方法是一种全新的模型辨识方法，它从牛顿迭代法发展而来，反复利用全局数据直到得到理想结果，可以充分发掘采样数据反映对象动态特性的潜力。而普通的递推最小二乘法则只能局部利用数据，能利用的信息量有限。这种方法可进行闭环辨识，无需外加激励信号，适

用于多输入多输出对象的在线辨识。本文提出的这种系统辨识方法，经过一系列的计算和验证，被证明是正确高效并且易于实现的。

(3) 本文以华中科技大学周怀春教授提出的采用辐射能反馈信号的火电机组负荷控制系统的思想为基础，对其中两个过程进行了系统辨识。一个为单输入单输出的给粉指令-辐射能过程，一个为从给粉指令和调门开度到机组功率和蒸汽压力这样一个双输入双输出的过程。这两个过程是前述火电机组负荷控制系统的被控对象组成部分。要实现这个控制系统，首先就要辨识出其被控对象的数学模型。本文对每一个过程都使用了多个模型结构进行试验，使用不同工况下的现场数据，并且每次计算过程都经过了自身辨识、交叉校验、求算术平均值、统一模型再校验这样几个步骤。通过比较和筛选，最后分别得到了最佳的模型。计算结果是令人满意的，对现场数据的预测输出拟合得非常好，可以认为这些模型是能够反映过程的本身特性的。

(4) 经过对最后得到的模型进行阶跃响应试验，发现给粉指令-辐射能信号这一过程的响应时间要比给粉指令-机组功率和给粉指令-蒸汽压力这两个过程的响应时间都要短得多，这也为上文描述的采用辐射能反馈信号的火电机组负荷控制系统的思想提供了佐证，是实现这一系统的有力保障。

6.2 下一步工作的建议和构想

(1) 本文将火电机组运行过程简化为一个双输入双输出过程，并以辐射能反馈信号作为内负反馈回路。这样的过程比较简单，省略了一些影响因素，甚至在某些工况下，这些因素会占据主导地位。正是这些因素造成了辨识结果与实际数据有一定的误差，影响了辨识的精度。

因而，下一步的工作是将火电机组的运行过程作一个更加复杂的假设，多考虑一些影响因素，构造一个多输入多输出系统，其输入输出项目

等于或者多于三个。在这样的系统中，会对过程产生影响的因素大部分被考虑进来，这样再进行多输入多输出系统辨识的时候就会得到更精确的结果。

(2) 本文提出的算法在进行系统辨识时，多使用机组负荷变化不太大的工况（一般不超过 10%）下取得的数据。如果现场数据的工况变化比较大，本算法的辨识结果就有一定的误差。因而下一步的工作还应该包括改进辨识算法，使之在进行系统辨识时，当遇到工况变化比较大的现场数据时也能够得到比较精确的结果。

(3) 本文进行辨识计算是基于采用辐射能反馈信号的火电机组负荷控制系统的思想而进行的，同时从一个侧面证明了这种思想的正确性。因而，在辨识出了这个系统的被控对象的模型参数后，就要进行控制系统仿真研究和现场试验，以实现采用辐射能反馈信号的火电机组负荷的优化控制。

致 谢

本文是在导师周怀春教授的悉心教诲和指导下完成的。在此，衷心的感谢我的导师周怀春教授。三年来，周老师一直关心我的学习和研究情况，积极的为我创造条件，不断的支持鼓励我，让我有足够的信心和勇气去面对学习中的各种困难。周老师严谨的科学作风、勤勤恳恳的工作态度、谦虚朴实的为人、渊博的知识以及勇于开拓创新的精神和能力，都给我留下了深刻的印象并深深地感染和教育了我，这些都将是今后学习和生活的楷模。在此我对我的导师周怀春老师致以衷心的感谢和最诚挚的敬意，祝愿周老师家庭幸福，工作顺利，事业蒸蒸日上，并祝好人一生平安。

感谢於正前工程师对我的关心。每次我在外地出差时，他都十分照顾我，无论在生活上还是在技术上都提供了很大的帮助。感谢姜志伟硕士，我们在一同出差的日子里互相帮助相互促进，给我留下了很难忘的印象。

感谢师兄韩曙东博士、娄春博士和师姐张师帅博士，三位博士师兄师姐在本科研究领域具有深厚的功底，与他们一起探讨问题总是充满了乐趣，感谢三位博士师兄对我的照顾和帮助。还要感谢已经毕业的师兄游君硕士、贾涛硕士，师姐侯玉波硕士，衷心祝愿他们今后事业有成，家庭幸福。衷心的感谢同课题组的李言钦博士、何其伟硕士、吴锐硕士、朱志武硕士，王大伟硕士、李莉硕士、刘浩硕士、吴峰硕士、杨超硕士，以及师弟程强硕士、阮刚硕士、方庆艳硕士和师妹曾佳硕士等等，他们共同营造了一个团结向上的科研氛围，和他们相处的时光是快乐而难忘的。希望他们今后继续努力工作，在周老师的带领下，使课题组日益壮大，欣欣向荣，并衷心祝愿他们在今后的道路上事业有成，生活安康。

最后感谢各位评委老师。

李锦

2003年4月于华工园

参考文献

- [1]韩才元、徐明厚、周怀春、邱建荣著, 煤粉燃烧, 北京: 科学出版社, 2001
- [2]钱学森等著, 工程控制论(修订版)上册, 科学出版社, 1980
- [3]王寿云等编著, 系统工程名词浅释, 科学出版社, 1982
- [4]Zedeh, L.A., "From Circuit Theory to System Theory" Proc.TRE, Vol.50
- [5]A. S. Brower, S. H. Javid, A. S. Patel, T. D. Younkings and G. H. Quentin, Application of A Coordinated Control Design Method to A Coal Gasification Plant, Modeling & Control of Electric Power Plants, Como, Italy, 1983
- [6]Ping Yang and B. W. Hogg, Continuous-time Generalised predictive Control of A Boiler Model, Control of Power Plants and Power Systems, Munich, Germany, 1992, 93-96
- [7]方崇智、萧德云编著, 过程辨识, 清华大学出版社, 1987
- [8]胡勇、巨林仓等, 电厂热工过程 ARMA 多参数辨识模型及应用, 热力发电, 2000.2: 37-39
- [9]唐巍、郭镇明等, 混沌优化用于电力系统静态负荷模型参数辨识, 电力系统自动化, 2000.7: 27-30
- [10]王作宏、王永骥, 基于多层前向神经网络的电加热炉系统辨识, 武汉科技学院学报, Vol.14, No.2, 2001: 19-25
- [11]夏超英, 基于复数模型描述的感应电动机参数辨识算法, 中国电机工程学报, Vol.20, No.12, 2000: 1-6
- [12]林景栋、曹长修等, 基于拓扑辨识的配电网故障定位算法, 重庆大学学报(自然科学版), Vol.24, No.5, 2001: 51-54
- [13]徐文、王大忠等, 模糊辨识在电力设备故障诊断中的应用, 电力系统自动化, Vol.20, No.3, 1996: 45-49
- [14]田松礼、陈亚民, 遗传算法用于电力系统频率特性参数辨识, 福州大学学报(自然科学版), Vol.28, No.4, 2000: 36-39

- [15]周怀春, 娄新生, 尹鹤龄等, 采用辐射能信号的电站锅炉燃料调节对象建模及控制仿真研究, 中国电机工程学报, Vol.16, No.4, 1996: 226-229
- [16]H. Nakamura Y. Toyota M. Kushihashi and M. Uchida, Optimal Control of Thermal Power Plants, Journal of Dynamic Systems, Measurement and Control, Transactions of the ASME, 1989,111:511-520
- [17]Haipeng Zhao, Wei Li, Cyurs Taft and Joseph Bentsman, Robust Controller Design for Simultaneous Control of Throttle Pressure and Megawatt Output in A Power Plant Unit, International Journal of Robust and Nonlinear Control, 1991, 9: 425-446
- [18]沈炯等, 火电单元机组负荷最优预见控制系统仿真研究, 中国电机工程学报, Vol.19, No.3, 1999: 14-17
- [19]张师帅, 周怀春等, 采用辐射能反馈信号的火电单元机组负荷控制系统仿真研究, 中国电机工程学报, Vol.21, No.2, 2001: 85-88
- [20]王永初, 闭环在线辨识对象模型的一种实用方法——(I) 初始模型的辨识, 自动化与仪器仪表, No.3, 1996: 10-12
- [21]张颖, 冯纯伯, 应用最小二乘法辨识闭环系统, 自动化学报, Vol.22, No.4, 1996: 452-455
- [22]史忠科, 模型在线辨识方法及其应用, 控制理论与应用, Vol.12, No.6, 1995: 787-791
- [23]戴义平, 俞茂铮, 刘炯, 王勇, 基于相关分析的汽轮机调节系统参数辨识研究, 汽轮机技术, No.5, 2001: 276-279, 297
- [24]李遵基主编, 热工自动控制系统, 中国电力出版社, 1997
- [25]翁思义, 自动控制系统计算机仿真与辅助设计, 西安交通大学出版社, 1987
- [26]章臣樾. 锅炉动态特性及其数学模型, 水利电力出版社, 1986
- [27]Rolf Isermann, Fault Diagnosis of Machines Via. Parameter Estimation

- and Knowledge Processing, Tutorial Paper, Automatical, 1993
- [28]Akaike, H., New Look at the Statistical Identification Mode, IEEE Trans. Auto. Control, Vol.19, 1974: 716-723
- [29]Schwarz, G., Estimating the Dimension of a Model, Ann. statist, Vol.6, 1978: 461-464
- [30]田铮, 动态数据处理的理论与方法——时间序列分析, 西北工业大学出版社, 1995
- [31]张家琛, 火电厂仿真, 水利电力出版社, 1994
- [32]张箭等, 基于 BP 网络的锅炉过热汽温系统动态特性的辨识, 电工技术学报, 1997
- [33]王耀清等, 燃烧控制系统最佳风/煤比曲线在线自学习算法, 华中理工大学学报, 1996, No.3: 32-35
- [34]J. Darlington, C.C. Pantelides, B. Rustem. B. A. Tanyi, An Algorithm For Constrained Nonlinear Optimization Under Uncertainty, Automatical Vol.35, 1999: 217-228
- [35]G. Baier, V. Schule, A. Vogel, Non-Nernstian Zirconia Sensors for Combustion Control, Appl. Phys.Vol. A57, 1993: 51-56
- [36]David St. John and Scott Samuelsen, Robust Optimal Control of a Natural Gas-Fired Burner for the Control of Oxides of Nitrogen (NO), Combust. Sic. and Tech, Vol. 128, 1997: 1-21
- [37]Liang Xingang, Guo Zengyuan, Liu Yunlong, et al, Ground Simulation of the Flow and Heat Transfer Under Microgravity, Sci. China(Series A), Vol.27, 1997(11): 1009-1013
- [38]Powe R. E., Carley. C. T., Bishop E. H., Free Convective Pattern in Cylindrical Annuli, ASME J Heat Transfer, Vol.91, 1969:310-314
- [39]刘浩等, 模糊数学在状态估计不良数据检测辨识中的应用, 电力系统及其自动化学报, 1996 (3)

- [40]方舒燕、杨乃贵、连世元等, 负荷模型对电力系统暂态稳定计算的影响, 电力系统自动化, Vol.23, No.19, 1999
- [41]鞠平, 电力系统负荷建模理论与实践, 电力系统自动化, Vol.23, No.19, 1999
- [42]沈善德, 朱守真, 罗骏等, 快速 BP 网络在负荷动态建模中的应用, 电力系统自动化, Vol.23, No.19, 1999
- [43]张琦, 韩祯祥, 文福栓, 进化规划方法在电力系统表态负荷模型参数辨识中的应用, Vol.14, No.4, 1997
- [44]唐巍, 李殿璞, 陈学允, 模型电力系统稳定器的混沌优化方法, 电力系统自动化, Vol.24, No.9, 2000
- [45]Zhu Jun, David L. Lubkeman, Adly A. Girgis, Automated Fault Location and Diagnosis on Electric Power Distribution Feeders, IEEE Trans. On PWRD, Vol.12, No.2,1997: 801-809
- [46]杜红卫, 孙雅明, 刘弘靖等, 基于遗传算法的配电网故障定位和隔离, 电网技术, Vol.24, No.5, 2000: 52-55
- [47]Lee Ying Chin, Hwang Chyi, Shih Yen Ping, A Combined Approach to Fuzzy Model Identification, IEEE Transactions on System, Man. and Cybernetics, Vol.24, No.5,1994
- [48]叶海文等, 燃煤循环流化床锅炉模型化的人工神经网络方法, 清华大学学报, Vol.37, No.2, 1997: 19-23
- [49]尹尔为等, 锅炉燃烧系统的模糊控制研究, 洛阳工学院学报, Vol.15, No.2, 1994
- [50]任雪梅等, 基于神经网络非线性系统辨识和控制的研究, 控制理论与应用, Vol.12, No.2, 1995: 147-153
- [51]韩志明, 李政, 倪维斗, Shell 气化炉的动态建模和仿真, 清华大学学报, Vol.39, No.3, 1999: 111-114
- [52]Schoen P., Dynamic Modeling and Control of Integrated Coal Gasification

华中科技大学硕士学位论文

Combined Cycle Units, Delft: Delft University of Technology, 1993

[53]Walve K., Modeling of Power System Components at Severe Disturbances, CIRGE Paper, 1986: 18-38

[54]鞠平, 马大强, 电力系统负荷建模, 北京: 水利电力出版社, 1995

[55]沈善德, 电力系统辨识, 北京: 清华大学出版社, 1993

攻读学位期间发表论文目录

- 1、李锦，周怀春，於正前，一种多输入多输出系统模型辨识方法及其应用，已被自动化与仪器仪表收稿