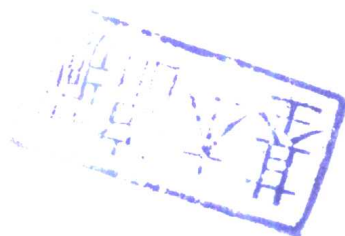


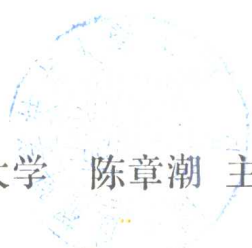
919616

高等学校教材



地区电网规划与监控

上海交通大学 陈章潮 主编



高等学校教材

地区电网规划与监控

上海交通大学 陈章潮 主编

水利电力出版社

高等学校教材
地区电网规划与监控
上海交通大学 陈章潮 主编

*

水利电力出版社出版
(北京三里河路6号)
新华书店北京发行所发行·各地新华书店经售
水利电力出版社印刷厂印刷

*

787×1092毫米 16开本 8.5印张 185千字
1990年10月第一版 1990年10月北京第一次印刷
印数 0001—4610册
ISBN7-120-01138-3/TM·333
定价1.80元

内 容 提 要

全书共五章。第一章介绍了适用于地区电网的几种负荷预测方法；第二章阐述了地区电网规划的内容和优化方法；第三章叙述了地区电网的无功功率电源优化配置；第四章介绍了地区电网计算机实时监控功能和软、硬件构成；第五章讨论了变电所和地区电网的无功电压优控原理及方法。

本书既可作为电力系统及其自动化专业本科生的专业选修课教材，又可作为电力系统工程技术人员的参考书。

前 言

目前的电力系统及其自动化专业教材和参考书中很少专门讨论地区电网或配电网的规划和运行问题，而电力工业中却需要有较多的工程技术人员从事这方面的工作。写作本书的目的就在于此。

本书是根据高等学校电力系统及其自动化专业学生的《地区电网规划与监控》选修课大纲要求编写的。全书共分五章，主要内容包括两大部分，第一部分是地区电网的规划，第二部分是地区电网的运行监视与控制。

本书由上海交通大学廖培鸿（第三章和第五章）和陈章潮（第一章、第二章和第四章）执笔编写，并由陈章潮任全书主编。全书由王平洋主审。

本教材在试用过程中曾得到能源部电力科学研究院王平洋同志和能源部电力科学技术情报研究所黄晞同志的指教及帮助，在此深表谢意。

由于编者水平有限，书中错误之处，恳请读者指正。

编 者

目 录

前言	
绪论	1
第一章 电力系统负荷预测	3
第一节 概述	3
第二节 时间序列预测法	4
第三节 经济模型预测法	16
第四节 地区小型水电站的发电量预测	24
第五节 多重相关预测法	27
第六节 网格预测法	31
第七节 预测误差分析	32
第二章 地区电网规划	34
第一节 概述	34
第二节 地区电网电压等级和接线方式的确定	35
第三节 地区电网规划的优化概念及其描述方法	37
第四节 变电所最优供电区的确定	41
第五节 最优配电系统规划的制定	44
第三章 地区电网无功电源规划	53
第一节 概述	53
第二节 目标函数	54
第三节 被控制量与控制量的约束方程	58
第四节 综合灵敏度矩阵	59
第五节 网络参数修正	62
第六节 计及二阶项的快速潮流计算	64
第七节 带有上下界控制变量的线性规划算法	68
第八节 整数规划算法	73
第九节 动态规划算法	77
第四章 地区电网计算机实时监控	83
第一节 概述	83
第二节 SCADA系统的功能	86
第三节 地区电网实时监控系统中的数据处理	90
第四节 地区电网监控系统的硬件结构	94
第五节 地区电网监控系统的软件结构	98
第五章 地区电网无功与电压最优控制	102
第一节 概述	102

第二节 电压与无功功率控制的几个基本概念	103
第三节 变电所的无功-电压综合控制	104
第四节 电力系统的无功-电压最优控制	109
附录 I 负荷分配系数 LDF 的计算	124
附录 II 线路损耗计算	126
参考文献	127

绪 论

地区电网经历过和主网相同的发展过程,当前同样也面临着如何做好发展规划、如何提高管理水平和运行水平问题,也就是如何积极地把新技术应用于地区电网,以提高地区电网建设和运行的经济效益。作为整个电力系统结构中一部分的地区电网,不论在网络建设还是运行管理方面都具有它本身的特点。为此,着重讨论这方面的问题是十分必要的。

电力系统负荷预测是进行电网规划的基础工作,其方法很多,有短期、中期和长期三类。每一类型中根据预测对象和预测目的又可选择各种建模方法。本书第一章所介绍的几种方法都适用于地区电网负荷性质的特点。时间序列法,从原理上讲,既能预测短期负荷,也能预测长期负荷,既能预测确定性的负荷,也能预测随机性的负荷。本书中着重强调时间序列法是处理随机负荷的好方法,往往和其它方法结合使用。经济预测模型的物理概念很清楚。因为电力负荷与国民经济发展、人民文化物质生活水平提高息息相关,只要搞清楚用电需要情况,电量的预测肯定可以比较“准”。这种经济模型预测方法对中小城市和地区是最合适的。网格法最适用于城市,包括低压配电网,往往根据负荷密度的大小来作规划,国外也有很多人采用。相关法可用来进行24小时调度预测。对于山区电网,很需要考虑小水电站对电网运行的影响,但是要确切掌握小水电站群的发电能力是不容易的,主要是由于降雨量难以确切掌握。国外生活用电中空调、冷藏、取暖等用电负荷激增,成为负荷预测中的重要项目。我国地区负荷中生活用电负荷部分尚不突出,但发展的趋势是值得注意的。

第二章讨论了地区电网的有功规划,简要地介绍了地区电网规划的主要任务和步骤。网络的电压等级是技术经济综合比较的结果,在作扩建规划时,必须考虑到原有电网电压等级的现状。从负荷增长的趋势、减少有色金属消耗和降低线损的角度看,提高配电网的电压水平是有益的。电网的接线方式取决于负荷的重要性。本章重点讨论优化理论在地区电网规划中的应用。建立目标函数和约束方程是进行优化计算的第一步。求解方程的方法很多,本章着重介绍线性运输法,这种方法概念明确,计算简便。此外还介绍了一种用数值解方法求得变电所的最优供电区和线路的最优供电区。

地区电网的无功电源合理配置也是地区电网规划的重要内容,并直接关系到电能质量和电网建设及运行的经济效益。第三章中首先讨论了为达到各种要求(如以网损最小、以补偿电容最小、以补偿效果最好和综合经济效益最好等)而建立相应的目标函数。达到这些目标的前提就是必须满足(保证)电网的状态参量(各母线的运行电压和各支路的潮流)不能超过容许的范围。本书在讨论时,以电压和潮流为被控制量,以移相电容量和带负荷调节变压器的分接头为控制量。为了求得被控量(或其增量)和控制量(或其增量)之间一一对应的数值关系,书中详细分析了综合灵敏度矩阵表达式,并推导出灵敏度矩阵的元素计算方程,这样就能方便地求出为了校正一定的电压偏差所需的电容增量或变压器

分接头增量。电容量和变压器分接头的改变,从网络结构方面看,也就是网络参数发生了变化,在作网络计算时必须注意到这一点。和做有功功率优化规划一样,无功电源优化规划也可以用各种数学规划来实现,本书只介绍带上、下限的线性规划法,这种方法计算比较简单,但有时收敛性欠佳,难于找到最优解。整数规划法虽然简化了计算,但是在变量比较多的情况下,计算也会变得复杂。以上两种方法存在一个共同缺点,就是不能描述变量的动态过程。动态规划法正好弥补了以上方法的这种不足,但又暴露了其固有缺陷,即方程的维数随着变量的增加迅速增大,计算时间和内存空间的占用也大为增加。这里也还是借助于灵敏度矩阵妥善地解决了这个问题。

第四章主要介绍了地区电网计算机实时监控的功能,功能的多少决定于地区电网的容量大小,也就是取决于它在系统中的地位。数据采集和处理是实现安全监视的基础。本书以实际监控系统为例,介绍了常用的数据处理方法和程序框图。硬件系统是以IBM-PC/XT为中心,以TP-86A为前置机,配1:4的远动装置、大屏幕显示和打印机等,并以Intel8251A为芯片的接口较好地实现了中断信息的处理,提高了分辨事故的能力。软件系统以DOS为环境,以BASIC高级语言和8086汇编语言为工具,详细介绍了应用软件功能和管理软件。数据库设计是十分重要的工作,本书把数据分为静态和动态的两种,分别加以处理和存放,便于存取又节省占用空间。模块式的应用软件结构,保证了使用的灵活性。

第五章是控制方面的内容。首先讨论了以一个实际变电所为对象的优化控制系统,对一次系统作了理论分析,并介绍了优化控制系统的硬件和软件结构。这种优化是局部的优化,不是整体优化。本章最后讨论了地区电网的整体优化控制原理和方法,具体探讨了无功电压静特性的辨识方法,用图论原理分析了电网结构,只有在确定当今网络结构基础上才能进行计算和控制,以网损最小为优化控制的目标函数。

本课程是电力系统及其自动化专业的选修课程,是一门用优化理论和计算机来解决地区电网规划和控制中的技术问题的专业课。其主要目的和任务是:通过学习本课程,扩大学生知识面,开阔思路,使他们懂得地区电网和高压大系统一样有大量的技术问题有待于人们去研究解决。将本课程和毕业设计等教学环节相配合,可使学生在解决实际工程问题方面的能力得到训练,为毕业后从事电网规划、科研和管理工作打下基础。

本课程应在学完《电力系统稳态分析》专业课程、学过数学规划基本知识的基础上进行教学。

本课程为30(或40)学时,分配如下:

绪论	1(1)学时
电力系统负荷预测	6(8)学时
地区电网规划	7(8)学时
地区电网无功电源规划	7(10)学时
地区电网计算时实时监控	5(7)学时
地区电网无功电压最优控制	4(6)学时

第一章 电力系统负荷预测

第一节 概 述

电力系统负荷预测是电力建设、调度的依据。例如，五年计划期间全国将增加多少发电容量、增建多少公里输电线路都必须与这期间的国民经济发展及人民生活水平的提高相适应，否则，将由于电力的不足影响经济的发展，或过剩的电力设备不能发挥作用，使大量的资金积压而无法获得效益。电力系统的调度部门更是如此，电力作为一种商品，要求供需随时平衡，负荷预测就相当于市场预测。而且“电”是一种特殊商品，发、供、用同时完成，无法象普通商品那样可以储存。一个大的电力系统有成百上千台发电机组，要决定哪些发电机该停运，运转的发电机各带多少负荷最为合适等等，都和当时的系统负荷大小密切相关。

电力负荷预测不仅体现在经济效益上，也是实现电网调度自动化，进行电网状态估计、安全分析功能的基础工作。

不难理解，如果电力系统负荷的预测值和实际值相差过大，就失去了预测的意义，但要做到“准”却有一定难度，原因是多方面的。其一，影响电量变化的因素较复杂，比如经济政策的调整，轻重工业比例变动，乡镇工业发展等等，都直接影响电量消耗；气象变化与用电量之间的关系很密切，在规划中还必须考虑天气负荷，在运行中要估计晴、雨或温度的突变对负荷的影响，其定量分析却不容易。其二，我们的基础工作做得不够好，对原始数据的收集、整理和积累欠重视，有关历史资料残缺不全，甚至于有“伪”数据，这样就难以反映电力负荷变化的历史本来面貌。

随着对电力负荷预测准确度要求的提高和计算工具的发展，预测方法也有很大的发展，不再只满足于单凭经验的预测方法。按预测时间的长短可分为：①短期预测，包括几十分钟的超短期预测和(24~48)h的短期预测，其具体方法有卡尔曼滤波法、相关法和时间序列法等；②中期预测（约1~5年）；③长期预测（6~15年或更长一些时间）。后两种可用时间序列法、经济模型和网格法等。以上各种方法的计算过程全由计算机来完成，除超短期预测因对运算时间约束较严格而采用在线计算外，其余都用于离线计算，对用机时间没有苛刻要求。

电力系统负荷预测是专门研究电力负荷的变化规律，以准确的统计数据和调查资料为依据，即从用电量的历史和现状出发，用科学的方法预测未来电力负荷的发展趋势和变化规律。在研究分析电力负荷的变化规律时，可采用各种不同的方法，只要它们都能满足预测准确、节省计算时间的要求。

这一章着重介绍时间序列法、相关法、经济模型法和网格法的基本原理及计算实例。

第二节 时间序列预测法

一、时间序列

时间序列可以用来描述具有确定性和随机性变化特征的事物。这种模型预测的准确性高,虽然计算较复杂,仍得到广泛的应用。

对于一个随机过程,认为它服从正态分布并符合一次马尔可夫过程,假定随机干扰量的期望值为零,则称此随机过程为平稳随机过程。实际工程上的数列往往不符合上述条件,大多为非平稳过程。不过大部分数列经过适当处理后,可以将非平稳数列变为平稳数列。下面先讨论处理时间序列用到的基本数学工具。

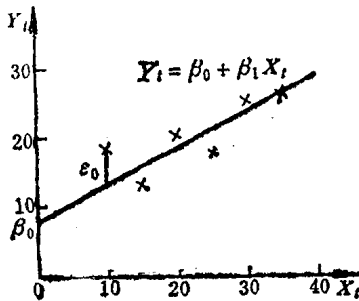


图 1-1 不同温度下的气体粘度

件,大多为非平稳过程。不过大部分数列经过适当处理后,可以将非平稳数列变为平稳数列。下面先讨论处理时间序列用到的基本数学工具。

1. 简单线性回归

假定观察值 Y_t 和 X_t 之间有依赖关系,比如在不同温度 $X_t (t=1, 2, 3, \dots, N)$ 和气体的粘度 Y_t 之间的关系,可以在 Y_t-X_t 座标平面上表示出来,如图1-1所示,然后用一阶线性回归方程来描述 Y_t 对 X_t 的关系,即

$$\dot{Y}_t = \beta_0 + \beta_1 \dot{X}_t + \varepsilon_t \quad (t=1, 2, \dots, N) \quad (1-1)$$

这是一条直线,它的截距为 β_0 ,斜率为 β_1 。认为各点至直线的偏离量 ε_t 的平方和为最小时,这条直线为这些点的最好拟合。根据这样的要求可以确定 β_0 和 β_1 的值为

$$\hat{\beta}_1 = \frac{\sum_{t=1}^N (Y_t - \bar{Y})(X_t - \bar{X})}{\sum_{t=1}^N (X_t - \bar{X})^2} \quad (1-2)$$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} \quad (1-3)$$

式中

$$\bar{X} = \frac{1}{N} \sum_{t=1}^N X_t, \quad \bar{Y} = \frac{1}{N} \sum_{t=1}^N Y_t$$

这样求得的 $\hat{\beta}_0$ 和 $\hat{\beta}_1$ 叫做 β_0 和 β_1 的最小二乘方估计。因为 ε_t 的平方和最小。符号“ $\wedge$ ”表示估计值,以区别于真值。

如果用 $Y_t = \dot{Y}_t - \bar{Y}$ 和 $X_t = \dot{X}_t - \bar{X}$ 分别代替(1-1)式中的 Y_t 和 X_t ,则新的一阶回归方程为

$$Y_t = \beta_1 X_t + \varepsilon_t \quad \left(\begin{array}{l} t=1, 2, \dots, N \\ \varepsilon_t \sim \text{NID}(0, \sigma_\varepsilon^2) \end{array} \right) \quad (1-4)$$

式中NID表示 ε_t 是正态独立分布。现在的 β_1 和 σ_ε^2 的最小二乘估计为

$$\beta_1 = \frac{\sum_{t=1}^N Y_t X_t}{\sum_{t=1}^N X_t^2} \quad (1-5)$$

$$\sigma_{\varepsilon}^2 = \frac{1}{N} \sum_{i=1}^N (Y_i - \beta_1 X_i)^2 \quad (1-6)$$

方程(1-4)指出了 Y_i 对 X_i 的依赖关系,它包括两部分,即相关部分 $\beta_1 X_i$ 和独立部分 ε_i 。 ε_i 是个未知量,除非 Y_i 和 X_i 都已知时,才有

$$\varepsilon_i = Y_i - \beta_1 X_i \quad (1-7)$$

当 Y_i 未知时, ε_i 是个随机变量,其平均值为零,方差为 σ_{ε}^2 ,并假定不同时间 t 的 ε_i 值为正态分布。

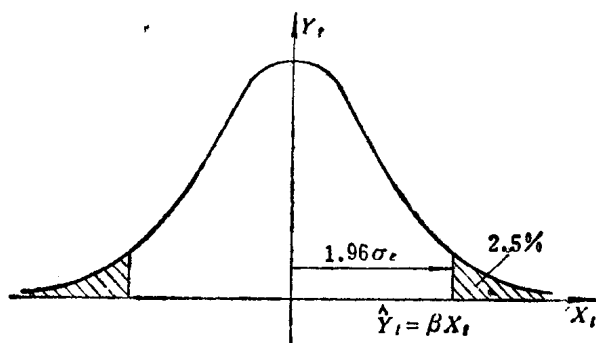


图 1-2 ε_i 的正态分布曲线

根据(1-4)式,可用已知的 X_i 来预测 Y_i ,即

$$\hat{Y}_i = \beta_1 X_i \quad (1-8)$$

但这并不是真值,真值应该为 $Y_i = \beta_1 X_i + \varepsilon_i = \hat{Y}_i + \varepsilon_i$ 。由于 ε_i 是正态分布,如图1-2所示,真值 Y_i 的95%有效出现在 $\hat{Y}_i - 1.96\sigma_{\varepsilon}$ 和 $\hat{Y}_i + 1.96\sigma_{\varepsilon}$ 之间。

下面用具体数据来举例说明上述参数的确定。

$$t = 1, 2, 3, 4, 5 \quad (N=5)$$

$$\dot{X}_i = 5, 6, 3, 2, 5$$

$$\dot{Y}_i = 7, 6, 5, 4, 6$$

则有

$$\bar{X} = \frac{1}{N} \sum_{i=1}^N \dot{X}_i = \frac{1}{5} (5+6+3+2+5) = 4.2$$

$$\bar{Y} = \frac{1}{N} \sum_{i=1}^N \dot{Y}_i = \frac{1}{5} (7+6+5+4+6) = 5.6$$

从观察值中减去平均值,即将 $X_i = \dot{X}_i - \bar{X}$ 代替 $\dot{X}_i$ 和 $Y_i = \dot{Y}_i - \bar{Y}$ 代替 $\dot{Y}_i$,得

$$X_i = 0.8, 1.8, -1.2, -2.2, 0.8$$

$$Y_i = 1.4, 0.4, -0.6, -1.6, 0.4$$

根据(1-5)和(1-6)式可得

$$\beta_1 = \frac{\sum_{i=1}^5 Y_i X_i}{\sum_{i=1}^5 X_i^2} = \frac{6.4}{10.8} = 0.59$$

$$\begin{aligned} \sigma_{\varepsilon}^2 &= \frac{1}{5} \sum_{i=1}^5 (Y_i - 0.59 X_i)^2 \\ &= \frac{1}{5} [0.928^2 + (-0.662)^2 + (0.102)^2 + (-0.302)^2 + (-0.072)^2] \\ &= 0.587 \end{aligned}$$

$$Y_i = \beta_1 X_i + \varepsilon_i = 0.59 X_i + \varepsilon_i \quad \left(\begin{array}{l} t=1, 2, 3, 4, 5 \\ \varepsilon_i \sim NID(0, 0.587) \end{array} \right)$$

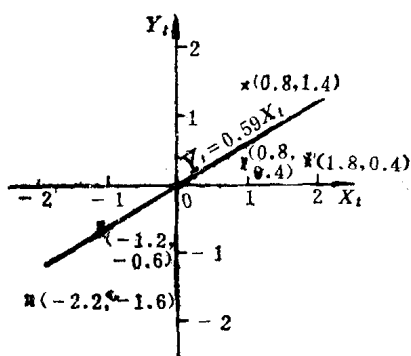


图 1-3 观察值的回归方程曲线

Y_t 的估计值为 $\hat{Y}_t = \beta_1 X_t = 0.59X_t$, 真值 Y_t 的 95% 概率可信度为

$$\begin{aligned}\hat{Y}_t \pm 1.96\sigma_e &= 0.59X_t \pm 1.96\sqrt{0.589} \\ &= 0.59X_t \pm 1.5,\end{aligned}$$

Y_t 对 X_t 的曲线如图 1-3 所示。

2. 多重回归

在上述简单一阶回归模型中, 认为气体粘度只和温度相关, 其实气体粘度还与浓度、蒸汽压力等

n 个变量 $X_{1t}, X_{2t}, \dots, X_{nt}$ 有关。和写方程 (1-4) 一样, 分别从各个观察值中减去其平均值后写成

$$Y_t = \beta_1 X_{1t} + \beta_2 X_{2t} + \dots + \beta_n X_{nt} + \varepsilon_t \quad \left(\begin{matrix} t=1, 2, \dots, N \\ \varepsilon_t \sim NID(0, \sigma^2) \end{matrix} \right) \quad (1-9)$$

其矢量形式是

$$\underset{\sim}{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_N \end{pmatrix} \quad \underset{\sim}{X} = \begin{pmatrix} X_{11} & X_{12} & \dots & X_{1N} \\ X_{21} & X_{22} & \dots & X_{2N} \\ \vdots & \vdots & & \vdots \\ X_{n1} & X_{n2} & \dots & X_{nN} \end{pmatrix} \quad \underset{\sim}{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} \quad (1-10)$$

按最小二乘估计参数

$$\underset{\sim}{\beta} = (\underset{\sim}{X}' \underset{\sim}{X})^{-1} \underset{\sim}{X}' \underset{\sim}{Y} \quad (1-11)$$

$$\sigma^2 = \frac{1}{N} \sum_{t=1}^N (Y_t - \beta_1 X_{1t} - \dots - \beta_n X_{nt})^2 \quad (1-12)$$

则 Y_t 的估计值为

$$\hat{Y}_t = \beta_1 X_{1t} + \beta_2 X_{2t} + \dots + \beta_n X_{nt}$$

真值的 95% 概率极限值是

$$\hat{Y}_t \pm 1.96\sigma_e \quad (1-13)$$

或

$$\beta_1 X_{1t} + \beta_2 X_{2t} + \dots + \beta_n X_{nt} \pm 1.96\sigma_e$$

3. 一阶自回归

用一阶回归描述 Y_t 和 X_t 之间关系的形式来描述变量自身的 X_t 对 X_{t-1} 、 X_{t-1} 对 X_{t-2} 等的关系, 称为一阶自回归。其形式为

$$X_t = \phi_1 X_{t-1} + a_t \quad (1-14)$$

从式中可看出 X_t 和它前面的量值 X_{t-1} 之间有密切依赖关系, X_{t-1} 值大, X_t 值也大; X_{t-1} 值小, 则 X_t 值也趋向小。式中的 ϕ_1 和 a_t 分别代替了 (1-4) 式中的 β_1 和 ε_t , 以区别于一阶回归方程, (1-14) 式称为一阶自回归, 记作 AR(1)。

这里必须指出, 如回归方程 (1-4) 一样, 假定不同时段 t 的噪声 a_t 是独立的, 即 a_{t-1} ,

$a_{t-2}, \dots$ 是彼此独立的干扰量, 而且服从正态分布, 即 $a_t \sim NID(0, \sigma_a^2)$ 。所以, 只要 ϕ_1 和 σ_a^2 被确定, (1-14) 式也就完全确定。其次是假定 a_t 与 X_{t-2}, X_{t-3} 等无关。这一点很容易理解, 因为 a_t 独立于 $a_{t-1}, a_{t-2}, \dots$, 这实际上就包括了第二假定。

AR(1) 模型的 ϕ_1 和 σ_a^2 参数估计为

$$\phi_1 = \frac{\sum_{t=2}^N X_t X_{t-1}}{\sum_{t=2}^N X_{t-1}^2} \quad (1-15)$$

$$\sigma_a^2 = \frac{1}{N-1} \sum_{t=2}^N (X_t - \phi_1 X_{t-1})^2 = \frac{1}{N-1} \sum_{t=2}^N a_t^2 \quad (1-16)$$

如果没有减去平均值, 则求 ϕ_1 的公式为

$$\phi_1 = \frac{\sum_{t=2}^N (X_t - \bar{X})(X_{t-1} - \bar{X})}{\sum_{t=2}^N (X_{t-1} - \bar{X})^2} \quad (1-17)$$

式中 ϕ_1 表示了 X_t 对 X_{t-1} 的依赖程度, 或者 X_t 对 X_{t-1} 的记忆能力强弱, ϕ_1 大即记忆性强, ϕ_1 小即记忆性弱, 但必须满足

$$|\phi_1| \leq 1 \quad (1-18)$$

充分性检验。用 (1-17) 式估计得到的参数进行计算是否合适, 还要作检验, 看看是否违背前述的假定, 即 a_t 对 a_{t-1} 和 $a_{t-2}, a_{t-3}, \dots$ 是否有依赖关系。检验方法是计算它们的自相关系数, 即

$$\hat{\rho}(a_t, a_{t-1}) = \frac{\sum_{t=2}^N a_t a_{t-1}}{\sum_{t=2}^N a_t^2} \quad (1-19)$$

和

$$\hat{\rho}(a_t, X_{t-1}) = \frac{\sum_{t=2}^N a_t X_{t-1}}{\sqrt{\left(\sum_{t=2}^N a_t^2\right) \left(\sum_{t=2}^N X_{t-1}^2\right)}} \quad (1-20)$$

若计算得到的 $\hat{\rho}(a_t, a_{t-1})$ 和 $\hat{\rho}(a_t, X_{t-1})$ 值都很小, 则它们之间是独立的, 说明得到的 ϕ_1 和 σ_a^2 二参数是合适的。于是就可用 AR(1) 模型进行预测。如果已知 $t-1$ 时刻的 X_{t-1} 值, 则 t 时刻的 X_t 一步预测值为

$$\hat{X}_{t-1}(1) = \phi_1 X_{t-1} \quad (1-21)$$

它与真值之间的偏差为

$$e_{t-1}(1) = X_t - \hat{X}_{t-1}(1) = a_t \quad (1-22)$$

预测值的 95% 概率可信度是

$$\hat{X}_{t-1} \pm 1.96 \sigma_a \quad (1-23)$$

即

$$\phi_1 X_{t-1} \pm 1.96 \sigma_a$$

4. N阶自回归

若经过(1-19)和(1-20)式的计算,发现自相关系数比较大,可以肯定相应的 σ_a^2 也大。这就说明用一阶模型来拟合实际数列是不充分的,必须提高模型的阶数,如AR(2), AR(3), ..., AR(N)。其方程式为

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_n X_{t-n} + a_t \quad (1-24)$$

可以用线性最小二乘法估计其参数。

$$\text{令 } \underset{\sim}{Y} = \begin{bmatrix} X_{n+1} \\ X_{n+2} \\ \vdots \\ X_N \end{bmatrix} \quad \underset{\sim}{X} = \begin{bmatrix} X_n & X_{n-1} & \cdots & X_1 \\ X_{n+1} & X_n & \cdots & X_2 \\ \vdots & \vdots & & \vdots \\ X_{N-n} & X_{N-n-1} & \cdots & X_{N-n-n} \end{bmatrix} \quad \underset{\sim}{\phi} = \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_n \end{bmatrix} \quad (1-25)$$

$$\text{则 } \underset{\sim}{\phi} = (\underset{\sim}{X}' \underset{\sim}{X})^{-1} \underset{\sim}{X}' \underset{\sim}{Y} \quad (1-26)$$

$$\sigma_a^2 = \frac{1}{N-n} \sum_{t=n+1}^N (X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \cdots - \phi_n X_{t-n})^2 \quad (1-27)$$

AR(N)的平稳性必要条件是

$$\phi_1 + \phi_2 + \cdots + \phi_n < 1 \quad (1-28)$$

X_t 序列的平稳性可通过其两相邻项的相关系数来衡量。 X_t 和 X_{t-k} 的自相关函数为

$$\hat{\rho}_k = \frac{E[(X_t - \bar{X})(X_{t-k} - \bar{X})]}{\sqrt{E[(X_t - \bar{X})^2]E[(X_{t-k} - \bar{X})^2]}} \quad (1-29)$$

$$\text{或 } \hat{\rho}_k = \frac{\sum_{t=1}^{n-k} (X_t - \bar{X})(X_{t+k} - \bar{X})}{\sum_{t=1}^n (X_t - \bar{X})^2} \quad (1-30)$$

随着 k 的增大, $\hat{\rho}_k$ 值下降快的为平稳序列, 如图1-4所示, 否则便是非平稳序列, 如图1-5所示。



图 1-4 平稳序列的自相关函数

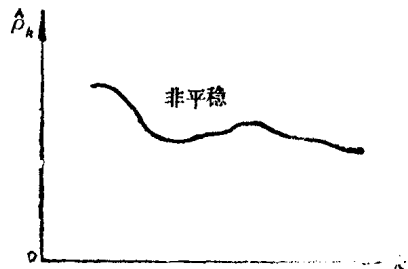


图 1-5 非平稳序列的自相关函数

随机模型必须具有平稳性要求, 但是齐次非平稳过程经过一次或数次差分后可以成为平稳性序列, 差分次数称齐次性的阶次。比如 X_t 是一阶齐次性非平稳性序列, 对它作一次差分的表达式是 $W_t = X_t - X_{t-1} = \Delta X_t$ 。符号 $ARI(n, p)$ 中的I和P分别表示差分和差

分运算的阶数。经过一次差分的自回归记作 $ARI(n, 1)$ 。二次差分表达式是 $W_t = \Delta^2 Y_t = \Delta Y_t - \Delta Y_{t-1}$ ，并记作 $ARI(n, 2)$ 。

5. 自回归与移动平均 $ARMA(n, m)$

$ARMA(n, m)$ 在特殊情况 $n \neq 0, m = 0$ 时成为 n 阶的自回归方程 $AR(n)$ ，而当 $n = 0, m \neq 0$ 时，则为移动平均方程 $MA(m)$ 。 $m = 1$ 称为一阶移动平均方程，即

$$X_t = a_t - \theta_1 a_{t-1} \quad (1-31)$$

X_t 对 a_{t-1} 有记忆作用。

$ARMA(n, m)$ 模型是包含 n 个 ϕ 和 m 个 θ 参数的非线性方程，即

$$X_t - \phi_1 X_{t-1} - \dots - \phi_n X_{t-n} = a_t - \theta_1 a_{t-1} - \dots - \theta_m a_{t-m} \quad (1-32)$$

或写成
$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_n X_{t-n} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_m a_{t-m}$$

 $(n > m, a_t \sim NID(0, \sigma_a^2)) \quad (1-33)$

从方程 (1-33) 可以清楚看到， $ARMA(n, m)$ 模型也有“记忆”功能，或者说具有动态特性，干扰量 a_t 对系统的影响不限于 t 时刻，而是一直延续到 m 个时间单位。

因为 $ARMA(n, m)$ 是非线性方程，所以要用非线性最小二乘来识别参数。从某个初值开始经过逐步迭代求得 a_t ，要求使其平方和达到最小。方程的阶数 n 和 m 是根据计算结果并作检验后加以确定。比如用非线性最小二乘估计求 $ARMA(2, 1)$ 模型的参数 ϕ_1 、 ϕ_2 和 θ_1 ，若其中 θ_1 值很小，那么有理由认为 $AR(2)$ 模型是合适的。即使用 $AR(2)$ 模型拟合结果，偏差量的平方和略大于用 $ARMA(2, 1)$ 模型拟合后产生的偏差量的平方和，故仍然可采用 $AR(2)$ 模型，并且 $AR(2)$ 模型的参数计算简单。

6. 检验

当 $ARMA(n, m)$ 模型阶数 n, m 决定后，应对模型的参数阶数作假设试验。利用标准统计学的 F 分布来作检验，看看增加或减少模型阶数对偏差影响的显著性，从而决定最合适的模型阶数。假定线性回归有 N 个观察值，有 γ 个参数。要减少 S 阶，则显著性检验准则为

$$F = \frac{A_1 - A_0}{S} \div \frac{A_0}{N - \gamma} \sim F(S, N - \gamma) \quad (1-34)$$

式中 $\gamma = n + m$ ， $A_0 = ARMA(n, m)$ 时偏差量 a_t 的平方和， $A_1 = ARMA(n, m)$ 降 S 阶后偏差量 a_t 的平方和， $F(S, N - \gamma)$ 表示自由度为 S 和 $N - \gamma$ 的 F 分布。

因为在静态情况下 $ARMA(n, m)$ 模型实际上也是线性回归，所以此检验准则也适用于 $ARMA(n, m)$ 模型。例如，对于 $ARMA(2n, 2n-1)$ 和 $ARMA(2n+2, 2n+1)$ 模型有 $\gamma = (2n+2) + (2n+1) = 4n+3$ ， $S = (4n+3) - 4(n-1) = 4$ ， $A_0 = ARMA(2n+2, 2n+1)$ 的 a_t 平方和， $A_1 = ARMA(2n, 2n-1)$ 的 a_t 平方和， N 为观察次数。则 F 准则为

$$F = \frac{A_1 - A_0}{4} \div \frac{A_0}{N - \gamma} = \frac{A_1 - A_0}{4} \div \frac{A_0}{N - (4n+3)}$$

$$\sim F(4, N - (4n+3)) \quad (1-35)$$

若 F 的计算值大于 $F(4, N - 4n - 3)$ 在分布表中 5% 的可信度，说明该模型从 $ARMA(2n,$

$2n-1$) 增加到 $ARMA(2n+2, 2n+1)$ 使偏差量减小是明显的, 即证明原来认为 $ARMA(2n, 2n-1)$ 的参数是充分的假定是不符合实际的, 应予以修正。反之, 若计算的 F 值小于查表值, 说明原来模型的参数不必作修正。

检验举例 假定发现 $ARMA(2, 1)$ 模型的参数 θ_1 很小, 且它的可信度范围包括零, 如果降阶使 $\theta_1 = 0$, 用 $AR(2)$ 模型对数列进行拟合, 结果有参数 $\phi_1 = 1.32$ 、 $\phi_2 = -0.63$, 它们的95%可信度范围不包括零。偏差量的平方和为41173.35, $ARMA(2, 1)$ 的偏差量平方和为4088.98, 于是有

$$A_1 = 41173.35, A_0 \approx 40788.00, \gamma = 2 + 1 + 1 = 4 (\text{包括一次平均值运算}), S = 3 - 2 = 1, N = 176$$

$$F = \frac{41173.35 - 40788}{1} \div \frac{4.0878}{176 - 4} = 1.62$$

查5%的 F 分布表得 $F_{0.05}(S, N) = F_{0.05}(1, \infty) = 3.84$ 。 F 的计算值和查表值作比较, 很明显 $F < F_{0.05}(1, \infty)$ 。同时 ϕ_2 参数的95%可信度范围不包括零, 因此, $AR(2)$ 是充分的。

二、用时间序列模型预测电力系统负荷实例^[1]

电力系统的负荷每时每刻都在变化, 可以把这种随时间变化的负荷数据看作一组时间序列。一般说, 它的变化过程不是一个平稳的随机过程, 因此, 不能简单地应用上述自回归方程来描述其变化规律。但是, 对历史的负荷数据作一番分析研究之后, 会发现负荷数据中存在按某种规律变化的成分, 也就是有确定性部分负荷, 而另一部分是捉摸不定的, 即所谓随机部分负荷。我们曾对我国某系统1980年的负荷作了分析, 发现有以下特点: ①月平均负荷增长率为1.5%~2.5%, 从月初到月终的日平均负荷也有一定增长, 但各月不等; ②每天(以小时为单位)、每周、每月(以日为单位)的负荷有明显的周期性; ③存在着相当比例的随机量。因此, 应该用确定的和随机的两种模型来描述全部负荷的变化过程, 即把负荷的历史记录数据分解成确定的和随机的两部分。所用的模型讨论如下。

1. 确定部分模型

电力负荷值的大小随时间而呈某种规律性地变化, 常常用以下几种模型来描述。

(1) 线性时序模型

$$Y_t = a + bt \quad (1-36)$$

式中 t ——时间;

Y_t —— t 时刻的时序值;

b —— Y 在 t 时刻的增长率。

(2) 指数时序模型

$$Y_t = e^{a+bt} \quad (1-37)$$

或把(1-37)式写成对数直线模型

$$\ln Y_t = a + bt \quad (1-38)$$

(3) 自回归模型

若把(1-36)式改写成

$$Y_t = a + bY_{t-1} \quad (1-39)$$