

张宝泉 刘庆东 张晓清 编著

计算机与环境

多因素分析



中国环境科学出版社

计算机与环境多因素分析

张宝泉 刘庆东 张晓清 编著

中国环境科学出版社

1993

(京)新登字089号

内 容 简 介

本书将环境、数学、计算机三者融为一体,系统地介绍了计算机在环境科学领域中的应用。阐述了多元统计分析的基本概念和方法,有助于读者提高运用计算机进行分析判断环境污染现状和制定防治环境污染决策的能力。

本书共分八章,主要内容是:介绍微机的操作系统,BASIC语言的基本知识和使用技巧;环境多元统计分析所必备的环境数学基础知识;多元线性回归分析、主成份分析、因子分析、判别分析、聚类分析等方法的数学原理和它们在环境体系中的应用实例。为环境评价、环境污染的控制与预报,为制定环境标准和排污收费办法提供科学方法和定量依据。

本书深入浅出,简明易懂,可供环保工作者及水文地质、气象、生物及有关工农业、卫生部门的科研、设计人员,大专院校师生、研究生等参考。

计算机与环境多因素分析

张宝泉 刘庆东 张晓清 编著

责任编辑 毛 平

•

中国环境科学出版社出版

北京崇文区北岗子街8号

北京市通县永乐印刷厂印刷

新华书店总店科技发行所发行 各地新华书店经售

•

1993年12月 第一版 开本 787×1092 1/32

1993年12月第一次印刷 印张 14 插页 2

印数 1-3000 字数 316千字

ISBN 7-80093-435-7/X·739

定价: 11.50元

前 言

环境科学是多学科、多门类综合性很强的边缘科学。许多环境问题是由于多因素综合造成的：必然因素、偶然因素，物理因素、化学因素、生物因素，以及人为因素等都在起作用。而且，环境系统中的各个因素之间又有着错综复杂的关系，环境系统中的各个方面发生变化将引起整个系统发生变化，环境系统是一个多维多元的体系。环保工作者要能够做到正确地评价环境质量，科学地制订环境标准，准确地提供环境污染预报，对环境进行现代化的管理，就需要把定性定量结合起来进行多监测点、多污染组分的测量，进行微观与宏观的综合研究。这就要求环保工作者必须很好地掌握多因素分析法与计算机的应用。本书的目的就是用简捷明了的方法阐述多元统计分析的基本概念和方法，帮助读者提高运用数学工具及计算机来分析判断环境污染现状和制定决策的能力。

本书的着眼点在于阐明多元统计分析推理的作用，各种方法的适用范围及其应用，而不是单纯地进行数学公式的讨论。在方法上，力求把各种方法内在的联系及各方法之间的联系，以及它们各自的特点说清楚，并通过一些数字、例子说明实施各种方法的主要步骤。在应用方面，尽量归纳和列举环境多因素分析的应用所涉及的环境问题，提出对多元分析结果进行环境解释的一些原则，并力求说清楚对各种不同的方法的计算结果进行环境解释时应着重注意的方面。

本书共分八章。第一章介绍微机的操作系统及BASIC语言程序的基本知识和使用技巧等内容。第二章介绍了环境多元统计分析所必备的环境数学基础知识,包括线性代数基础知识和概率统计基础知识等内容。第三、四、五、六、七、八章分别介绍了多元线性回归分析、主成份分析、因子分析、判别分析、聚类分析等方法的数学原理和它们在环境体系中的应用实例。读者可以根据书中所介绍的方法原理在其它领域作一些应用的尝试。

本书将环境、数学、计算机三者融为一体,介绍了一批具有实际应用价值的,简便易行的计算机程序,这样会有助于环保工作者进一步认识环境问题,提高工作效率。本书所给出的全部计算机程序,可在IBM-PC微型计算机及其兼容机上运行。本书配有计算机程序软盘,可供读者选用。

书中给出的部分实例是选自其它作者的报告,在此深表谢意。在本书的编写过程中得到了张忠占、毛平等同志的大力帮助和指导,另外北京纺织工程学院张宝杰进行了本书部分程序的调试工作,在此表示衷心的感谢。

由于我们对环境科学、计算机科学和数学的了解和研究还很不够,业务水平有限,书中错误与缺点难免,恳请读者、专家批评指正。

编者

1993年2月

目 录

第一章 微机系统基础与BASIC语言程序	(1)
第一节 磁盘操作系统及使用	(1)
一、DOS概述	(1)
二、DOS的启动	(4)
三、IBM-PC DOS命令格式	(8)
四、文件及其操作命令	(10)
五、常用磁盘操作命令	(22)
六、DOS的标准输入、输出的换向特性	(27)
七、修改文件属性	(33)
八、复制备份文件	(34)
九、修复文件命令	(36)
第二节 IBM-PC BASIC的使用方法	(37)
一、BASIC简介	(37)
二、BASIC的使用方法	(38)
三、BASIC的功能键与编辑键	(42)
四、BASIC编辑命令及文件管理命令	(44)
五、BASIC的运行、控制和输出打印	(50)
第三节 常见BASIC程序语句错误分析	(52)
一、关于BASIC指令错误	(53)
二、有关数组的错误	(57)
三、有关程序输入的操作错误	(60)
第四节 BASIC程序使用技巧	(62)
一、实现BASIC程序中PRINT和LPRINT自动转换	(62)

二、对BASIC程序的P参数文件解密.....	(68)
第五节 BASIC程序与其他软件的联合使用.....	(70)
一、在BASIC程序中调用各种软件和DOS命令.....	(70)
二、在BASIC中直接调用dBASE的数据文件.....	(74)
三、用BASIC程序文件调用由Wordstar建立的文件.....	(79)
第二章 数学基础.....	(82)
第一节 线性代数基础.....	(82)
一、行列式.....	(82)
二、矩阵及其运算.....	(85)
三、 n 维向量空间.....	(99)
四、线性代数方程组.....	(101)
第二节 环境统计的概率基础.....	(115)
一、随机事件及其概率.....	(115)
二、随机变量及其分布函数.....	(122)
三、随机变量的数学期望与方差.....	(132)
第三章 多元统计分析引论.....	(138)
第一节 引言.....	(138)
第二节 原始数组矩阵.....	(141)
一、原始数组矩阵的建立.....	(141)
二、建立原始数组矩阵的规则.....	(142)
第三节 原始数组矩阵的研究方法.....	(143)
一、对原始数组的研究方式.....	(144)
二、原始数据的简化.....	(145)
第四节 数据处理基础概念.....	(147)
一、原始数组的协方差矩阵.....	(148)
二、原始数组的相关矩阵.....	(149)
三、原始数据的标准化.....	(150)
第四章 多元线性回归分析及其应用.....	(152)

第一节 相关关系	(152)
一、概述	(152)
二、相关关系	(153)
第二节 多元线性回归方程的建立	(155)
一、最小二乘法	(155)
二、二元线性回归方程的建立	(156)
三、多元线性回归方程的建立	(157)
四、回归方程的检验	(160)
第三节 多元线性回归应用示例	(166)
一、示例	(166)
二、多元线性回归应用程序	(173)
第四节 逐步回归	(184)
一、最优回归方程	(184)
二、逐步回归方程的建立	(185)
三、逐步回归举例	(190)
四、多污染物控制与预测最优方程在计算机上的实现	(201)
第五章 环境主成份分析	(217)
第一节 主成份分析的基本原理	(218)
一、引例	(218)
二、主成份分析原理	(219)
第二节 主成份分析步骤	(223)
一、对原始数据进行标准化	(223)
二、由标准化后数据数组求协方差矩阵 Σ	(223)
三、计算特征值及特征向量	(224)
四、确定主成份	(225)
五、计算主成份得分	(225)
六、结论解释与推断	(225)
第三节 主成份分析的应用	(225)
一、在环境污染程度分析方面的应用	(225)

二、在优化布点方面的应用.....	(231)
三、在污染源识别方面的应用.....	(237)
第四节 主成份分析程序.....	(245)
一、程序结构图.....	(245)
二、程序清单.....	(246)
第六章 环境因子分析	(256)
第一节 引言.....	(256)
一、压缩原始数据.....	(257)
二、指示成因推理的方向.....	(258)
三、分解叠加的环境过程.....	(258)
第二节 因子分析的环境概念.....	(259)
一、R 型因子分析的环境概念.....	(259)
二、Q 型因子分析的环境概念.....	(260)
第三节 因子分析方法.....	(262)
一、因子分析的数学模型.....	(262)
二、因子分析步骤.....	(269)
三、因子分析与主成份分析关系.....	(278)
第四节 环境因子分析实例.....	(279)
一、大气污染源识别应用.....	(279)
二、主成份——因子分析应用.....	(294)
第五节 环境因子分析的计算机程序.....	(297)
一、程序框图.....	(297)
二、程序说明.....	(297)
三、程序清单.....	(300)
第七章 环境判别分析	(316)
第一节 Fisher 二类判别准则.....	(316)
一、原理.....	(316)
二、公式介绍.....	(317)
三、应用介绍.....	(322)

四、Fisher二级判别分析的程序结构图	(327)
第二节 Bayes判别准则	(328)
一、贝叶斯准则	(328)
二、多级判别计算步骤	(331)
三、贝叶斯判别效果的检验	(334)
四、实例分析介绍	(335)
五、贝叶斯多级判别分析程序结构图	(342)
第三节 逐步判别分析	(343)
一、逐步判别计算步骤	(343)
二、实例应用介绍	(346)
第八章 聚类分析及其应用	(359)
第一节 聚类分析的统计量及数据的均一化	(361)
一、度量相似程度的统计量	(361)
二、均匀化问题	(364)
第二节 系统聚类法	(366)
一、最短距离法	(366)
二、最长距离法	(378)
三、重心法	(380)
四、系统聚类分析程序	(381)
第三节 模糊聚类分析	(401)
一、建立模糊相似矩阵	(402)
二、创建模糊等价矩阵	(408)
三、进行聚类分析	(412)
四、模糊聚类分析的程序	(414)
附表	(425)
参考文献	(435)

第一章 微机系统基础与 BASIC语言程序

在环境科学的研究、管理及监督中将得到有关的大量数据,对于这些数据,只有采用正确的数学方法去分析和处理,才能探索到环境科学的规律,得到正确的结论。环境多元分析比起其它目前所采用的分析方法,更能揭示环境的本质规律。但由于多元分析计算复杂,数据多,计算量大,只有运用计算机才能使这一工作得以顺利进行。鉴于众多的环保工作者是非计算机专业的,而计算机的BASIC语言又易为人们学习和掌握,所以为了使广大环保工作人员能够较为顺利的运用IBM PC微型计算机进行数据分析处理,本章将简单介绍计算机DOS系统和BASIC语言的一些常用命令及其使用方法和技巧。

第一节 磁盘操作系统及使用

磁盘及其操作系统的使用是我们与微机打交道时首先遇到的问题。磁盘操作系统的正确使用,将会对机器的使用和软件的开发带来较大的影响。本节将主要讨论DOS基本命令及在DOS的使用过程中可能出现的一些情况和相应的处理方法。

一、DOS 概述

DOS是Disk Operation System的缩写。PC DOS是

美国Microsoft公司为IBM-PC计算机开发的磁盘操作系统，是帮助用户使用计算机、协助用户完成各种工作必不可少的一组程序。其功能主要是进行文件管理和设备管理。文件系统负责建立、删除、读写和检索各类文件，而I/O系统则负责驱动外围设备，如显示器、键盘、磁盘、打印机等。

(一) IBM-PC DOS的基本结构

PC DOS采用层次模块结构，它由三个层次模块和一个引导程序组成。这三个模块是：输入输出系统，文件系统(IBM DOS.COM)和命令处理程序(COMMAND.COM)。

1. 引导程序BOOT

它是一个很小的程序，它存放在DOS系统盘的首部，每次启动DOS时，它会首先自动装入内存，然后引导装入IBMBIO.COM、IBMDOS.COM两个文件，将控制权交给IBMDOS.COM。

2. 输入输出模块

由基本输入输出模块ROMBIOS和基本输入输出接口模块IBMBIO.COM两部分组成。该模块主要负责DOS与非盘外部设备的联系，在计算机内存和外部设备上读写数据，如完成对内存大小测试、芯片检查、键盘及打印机的初始化等。IBMBIO.COM以隐含文件的形式存放在DOS系统盘上。

3. 文件系统IBMDOS.COM

它是PC DOS的核心。它含有一个文件管理程序和磁盘读写及其它外设管理方面的功能子程序。它是以隐含文件的形式存放在DOS系统盘上。设计在DOS控制下运行的程序

可以使用这些文件管理程序和服务功能子程序。

4. 命令处理程序COMMAND.COM

它是存放在DOS系统盘上的基本文件，它负责接收用户由链盘输入的命令，然后运行相应的程序。

(二) CC-DOS的组成

CC-DOS即中文磁盘操作系统 (Chinese character Disk Operation System)，是在DOS的基础上扩充了汉字处理能力的操作系统。CC-DOS不要求对IBM-PC计算机做硬件改动，也不增加新的硬件就能处理文中信息，还能直接引用原DOS的程序和命令。

当用户要求处理中文时，必须调用CC-DOS，使机器在中文操作系统控制下工作。

CC-DOS由以下四个基本部分组成：

1. DOS系统文件

包含IBMDOS.COM, IBMBIO.COM和COMMAND.COM三个文件，其中前两个是隐含文件。

2. CC-DOS的主体文件

CC-DOS的主体文件有三个：FILE1.EXE、CCCC.EXE和CCLIB。其中：

FILE1.EXE文件是引导输入程序，负责给字库开辟内存区，并完成初始处理和模式切换。

CCCC.EXE文件负责装入CC-DOS和汉字库。

CCLIB文件是汉字字库程序，共有汉字8763个和619个图形符号，约占238k字节，是一种16×16点阵型的汉字字模。目前常用的还有一种专供打印用的24×24点阵汉字字模，当系统配有硬盘时，可将其存放在硬盘上，约占580k字

节, 文件名为: CLIB24, CLIB241, CLIB242。

3. 打印驱动程序

该驱动程序可确保打印机在打印汉字的过程中正常工作。不同的打印机需要不同的打印驱动程序, 同样的打印机在不同的微机系统中打印驱动程序也可能不一样。目前, 我们最常用的打印机有24针打印机和9针打印机。如D320.EXE, D32024.EXE等程序可驱动24针打印机, 打印 24×24 点阵汉字。ALL9P.EXE, NEW9P.EXE等是驱动9针打印机的, 打印 16×16 点阵汉字。

随着软件的不断开发, 已有更多的打印驱动程序供用户自行选择。

4. AUTOEXEC.BAT文件

这是个自动启动的批处理文件。DOS每次启动时, 命令处理程序就在DOS盘上寻找这个文件, 若找到, 则立即执行之。

二、DOS的启动

启动PC DOS就是用户启动计算机开始工作。也就是指把DOS系统由磁盘读出装入到计算机的内存, 并作好系统初始化工作, 供用户使用。

(一) 系统的冷启动

计算机接通电源后的开机启动称为冷启动, 其实现步骤为:

1. 把装有DOS系统的软磁盘插入驱动器A中, 并把驱动器的门关好。
2. 打开显示器电源。
3. 如果使用打印机, 则打开打印机电源。

4. 接通主机电源。这时计算机开始工作，首先进行自检，然后引导DOS，先从A盘上寻找DOS，此时A驱动器的指示灯不断闪烁，表示DOS正由盘上读出装入内存，稍等片刻后系统引导成功，此后屏幕将提示系统日期、时间，用户按提示的格式要求分别回答正确后，屏幕上显示DOS的版本及拷贝时间，最后显示DOS提示符“A>”：

The IBM Personal Computer DOS
Version 3.30(c) Copyright IBM Corp 1981, 1982,
1983, 1984

A>

屏幕上出现DOS提示符“A>”，表示现在计算机已处于DOS的控制下，等待用户输入DOS命令，开始进行需要的处理工作。

“A>”表示当前的驱动器为A驱动器冷启动至此完成。

(二) 系统的热启动

热启动是指系统已通电工作后的重新启动。通常，当用户发现计算机工作异常时采用这种启动方式，使计算机重新进入DOS状态，所以热启动又称系统复位。启动步骤为：

1. 将DOS软盘插入驱动器A中，把驱动器的门关好。
2. 按下Ctrl和Alt键的同时，按下Del键，然后一起松开，系统将重新启动。

热启动过程中显示的信息与需要的操作与冷启动时相同，但计算机不再进行硬件自检诊断。

此时热启动完毕。

有一点用户应注意：热启动将会使机器正在进行的工作

全部报废。

(三) 系统的硬盘启动

如果计算机带有硬盘机, 且其上装有DOS系统软件, 那么也可用硬磁盘启动DOS。

用硬盘启动DOS较软盘启动速度要快, 而且方便。

用硬盘启动DOS时, 应注意:

1. 硬盘中应有DOS文件。
2. 不要在A驱动器中插盘, 或要打开A驱动器的门。
3. 系统的冷、热启动的其余步骤同系统的软盘冷、热启动。
4. 此时DOS的提示符为“C>”。

(四) 输入日期

当DOS装入, 系统自检完毕后, 屏上立即出现提示:

Current date is 1-01-1980

Enter new date: -

输入当天的日期, 用“-”或“/”分隔月、日、年, 例如,
今天是1992年10月23日, 则键入:

10-23-1992 ↵或

10/23/1992 ↵

若系统提示时你认为无需输入日期, 则按↵键 (即Enter键)。

(五) 输入时间

输入一个有效日期后, 系统接着提示:

Current time is 0:01:37.84

Enter new time: -

输入当前时间, 用“:”号分隔小时、分、秒, 用“.”分隔秒、百分秒。

例如: 现在时间是8点52分, 则键入:

8:52:00↵

若你认为无需输入时间, 则按↵键。

此后, 系统进入DOS状态, 出现DOS提示符。

CC-DOS的启动与上述步骤完全相同, 仅将DOS系统盘换为CC-DOS系统盘即可。

(六) 指定当前驱动器

当前驱动器, 也叫缺省驱动器, 由DOS提示符中的字母指示, 如A>, B>等。改变当前驱动器是通过输入一个新指定的驱动器标识字符, 再按“:”号键和“↵”键来实现。如屏幕显示:

A> 说明当前驱动器为A

A>C:↵ 把当前驱动器转为C

C> 当前驱动器为C

(七) DOS常用的控制键

当操作人员在DOS的支持下输入程序或回答系统提问时, 都要用键盘向计算机输入信息。除空格、回车、大小写换挡和退格这些对任何软件都具有通用意义的键外, DOS为用户提供了一组对DOS系统起特殊作用的控制键, 使用它们可以对系统的运行进行一定程度的干预。

控制键多数由几个键同时动作组合而成。如Ctrl-C表示按住Ctrl键的同时按下C键。