



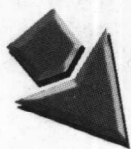
JISUANJIJISHU ZAISHENGMINGKEXUE
ZHONG DE YINGYONG

计算机技术 在生命科学中 的应用

郑秋生 编著

山东科学技术出版社
www.lkj.com.cn





JISUANJIJISHU ZAISHENGMINGKEXUE
ZHONG DE YINGYONG

计算机技术 在生命科学 中的应用

郑秋生 编著



山东科学技术出版社

图书在版编目 (CIP) 数据

计算机技术在生命科学中的应用/郑秋生编著. — 济南:
山东科学技术出版社, 2001
ISBN 7-5331-2816-8

I. 计… II. 郑… III. 计算机应用-生物学
IV. Q-39

中国版本图书馆 CIP 数据核字 (2000) 第 87112 号

计算机技术在生命科学中的应用

郑秋生 编著

出版者: 山东科学技术出版社

地址: 济南市玉函路 16 号

邮编: 250002 电话: (0531) 2065109

网址: www.lkj.com.cn

电子邮件: sdkj@jn-public.sd.cninfo.net

发行者: 山东科学技术出版社

地址: 济南市玉函路 16 号

邮编: 250002 电话: (0531) 2020432

印刷者: 山东人民印刷厂

地址: 泰安市灵山大街东首

邮编: 271000 电话: (0538) 6119354

开本: 787mm × 1092mm 1/16

印张: 16

字数: 353 千

版次: 2001 年 10 月第 1 版第 1 次印刷

ISBN 7-5331-2816-8

TP·72

定价: 26.00 元

序

《计算机技术在生命科学中的应用》一书随着 21 世纪曙光的来临问世了。该书出版可以说是生逢其时。人们公认,21 世纪初叶,信息科学和生命科学将是对人类社会进步产生最大促进作用的两大门类。本书恰好把信息科学所不可缺少的工具——电子计算机运用到生命科学的各个领域中去。可以预言,计算机生物学必将对生命本质的认识起着其他研究方法所不可替代的作用,同时,由于生命本质的复杂性和多样性,计算机生物学又将丰富和促进信息科学的发展。

作者得风气之先,在烟台大学开出了计算机技术在生命科学中的应用这门课,这门课既新颖又重要,至今在全国高等院校中也并不普遍。本书首先介绍基本数学模型,然后介绍计算机技术在生物化学、生理学、遗传学、生态学等领域的应用,还介绍了计算机技术在微生物学中的应用,这一内容在基本理论上以及发酵、制药、细胞工程等实际生产中都有参考价值。对生物学中普遍存在的图像处理和识别问题,书中也作了详细叙述。对生物统计学和生物学中常用的数学方法提供了实用程序,便于操作。应该再次强调的是,当把计算机这个现代最有力的研究和运算工具用之于生物学时,绝不仅仅是计算上的快速和简便,它必然会揭示用其他研究工具所不能揭示的重要规律。这是一个极新的领域,有志者在这一新天地中会大有作为,本书将为此提供入门钥匙。

郑荣梁

生物物理学教授、博士生导师

中国生物学会理事

前 言

生命科学是目前研究最活跃的学科之一,计算机在这一领域内越来越显示出它的重要性。曾有人预言,未来的生物学家应用计算机的能力将大大超过使用传统显微镜的能力。计算机计算技术在核酸序列分析、生物大分子空间结构预测、微生物发酵过程控制、酶反应动力学研究、生态学模拟等方面的应用已经表现出巨大的魅力。为此,作者在多年教学和科研的基础上编写了本书,旨在帮助读者对计算机在生命科学中的应用有一个较全面系统的认识。本书尽量采用简单的数学方法,不追求数学上严格的推导,读者利用现有的知识就能很快掌握本书的内容。书中模型多以专业著作的基本理论为前提,并附有大量的程序,用 Fortran 语言系统全面表述生物学中的一些基础知识和基本现象,并在计算机上进行模拟。

生物学者往往对复杂的数学推导过程一筹莫展,而数学家又对生物学的问题了解甚少,本书着重介绍计算机技术在生命科学研究过程中的应用,尽量简化数学过程,架起生命科学研究与数学模型的桥梁,使生物学者克服对数学常有的畏惧心理。

随着微型计算机的普及和运算速度的提高,计算机技术在生命研究中的作用越来越重要。书中所介绍的工作都能在 PC 微机上进行模拟,无需增加其他专门化的设备,便能很快学会用数学模型展示生命规律的方法,并把复杂的问题尽量简单化。很快掌握数学模型展示生命规律的方法,并在普通计算机上进行模拟,这正是生物学者所热切期望的。

本书从简单的模型入手,阐述计算机技术在生命科学中的应用。第一章简述计算机技术在生命科学中的应用;第二章介绍建立数学模型和处理模型的基本方法;第三章至第七章介绍计算机技术在生物化学、微生物学、遗传、生理学及生态学中的作用;第八章简单介绍图像处理的方法及其应用;第九章介



绍计算机在生物统计学中的应用;第十章介绍最优化方法在生命科学研究中的应用;第十一章介绍生物学中常用的数学方法,读者在编写程序时可直接调用所介绍的程序。

本书介绍的模型多是在简单条件下的模拟,并人为地将复杂的问题简单化。生命现象是复杂的,简单的模型组合成大型的复杂模型,去指导人们从事生产和研究活动,从而产生巨大的经济效益。当越来越多的生物学者用计算机手段研究生命现象时,随之而来的将是一场生物学研究上的革命。

本书出版得到于维纶学术出版基金的资助,在此谨致以诚挚的感谢。

编 者

目 录

第一章 绪论	1
一、概念模型与数学模型在生命科学研究中的重要性	1
二、计算机模拟与模型的建立步骤	3
三、微型计算机模型的建立	5
第二章 建立数学模型和处理模型的基本方法	6
一、微分方程模型	6
二、稳态假定模型	8
三、微分方程的数值解法	9
(一) 有限差分法	9
(二) 简单欧拉积分法	10
(三) 改良欧拉积分法	10
(四) 迭代二阶龙格 - 库塔法	11
第三章 计算机模拟技术在生物化学中的应用	12
一、二级反应动力学模型	12
二、化学平衡模型	16
三、酶底物相互作用的 Chance - Cleland 模型	18
四、自催化反应模型	20
五、pH 对酶活性的影响模型	22
六、别构酶的协同作用模型	25
七、温度对酶反应速度的影响模型	27
八、抑制剂对酶反应速度的影响模型	30
(一) 竞争性抑制模型	30
(二) 非竞争性抑制模型	32
(三) 反竞争性抑制模型	33
九、基因抑制对蛋白质合成的控制	33
十、单底物反应的一般质量作用方法	34
十一、单底物反应序列的速率定律方法	35
十二、Yates - Pardee 反馈控制模型	36
十三、磷酸果糖激酶的变构控制作用	39
十四、糖酵解的质量作用模型	41
第四章 计算机模拟技术在微生物学中的应用	46



一、Monod 模型	46
二、分批培养模型	50
三、连续培养或恒化器模拟	52
四、多种限制性营养物对微生物生长影响的模型	55
五、微生物对单一限制营养物竞争模型	58
六、微生物对多种限制营养物的竞争	60
七、毒剂的亚致死浓度对生长的抑制	61
八、两级营养物摄取模型	62
九、Droop 营养物限制模型	63
十、内部营养池对营养物摄取的控制	63
十一、多种营养物限制模型	64
十二、温度对发酵的影响模型	65
第五章 计算机模拟技术在遗传学中的应用	67
一、选择对显性突变的影响模型	67
二、平衡遗传负荷及平衡多态现象	69
三、平衡突变负荷	69
四、对伴性隐性基因的选择	70
五、对两位点隐性基因的选择	70
六、Wright 近亲繁殖模型	72
第六章 计算机模拟技术在生理学中的应用	73
一、简单扩散模型	73
二、渗透压模型	75
三、逆流扩散的室模型	76
四、主动运输模型	77
五、主动运输的简单处理方法	79
六、细胞外空间的测定	79
七、液流过程的室模型	80
八、心脏主动脉段的机械运动模型	81
九、碘代谢的室模型	82
十、二氧化碳浓度调节的简单模型	83
十一、甲状腺分泌的调节模型	85
十二、热中性点以上的体温调节模型	86
十三、热中性点以下的体温调节模型	88
第七章 计算机模拟技术在生态学中的应用	91
一、种群的内禀增长能力	91
二、逻辑斯蒂增长模型	92
三、环境容纳能力与群体密度的波动过程模型	94
四、标志重捕法模型	96



五、寿命表模型	99
六、狩猎模型	101
七、渔业模型	102
第八章 计算机图像处理技术在生命科学中的应用	104
一、图像处理的基础知识	104
(一) 图像表示方法	104
(二) 图像传输	105
二、血球自动分类处理	105
(一) 分割	106
(二) 特征抽取	106
(三) 分类	106
三、癌细胞识别系统	108
(一) 癌变细胞的特征	108
(二) 癌细胞的识别方法	108
(三) 形状识别	109
(四) 胞核光学密度	110
(五) 胞质面积	110
(六) 分级树状判别	110
四、染色体核型分类	110
(一) 染色体分析的要求	110
(二) 染色体核型分类	111
第九章 计算机在生物统计学中的应用	114
一、代表数和变异数	114
(一) 算术平均数 $\bar{X}$	114
(二) 标准差 S	114
(三) 标准误 S_x	115
二、 t 测验	116
(一) 配对资料比较的 t 测验	116
(二) 两组计量资料比较的 t 检验	118
三、方差分析	121
(一) 单因素方差分析	121
(二) 两因素方差分析	124
四、卡方检验	128
五、相关分析	135
六、回归分析	136
(一) 非线性回归分析	136
(二) 多元线性回归分析	140
(三) 多项式回归分析	145



第十章 最优化方法在生命科学研究中的应用	149
一、一维搜索方法	151
(一) 确定搜索区间	152
(二) 0.618 法——黄金分割法及其 FORTRAN 程序	154
(三) 二次插值法	159
二、无约束优化方法	164
(一) 梯度法(最速下降法)	164
(二) 拟牛顿法	167
三、约束最优化方法	170
(一) 内罚函数法(内点法)	170
(二) 扩展内罚函数法	171
(三) 二次扩展内罚函数法优化程序——CONSMIN3	173
(四) SUMT 方法使用指南	188
四、优化方法的应用实例	192
(一) 问题	192
(二) 模型建立	193
第十一章 生物学中常用的数学方法	196
一、插值与逼近	196
(一) 拉格朗日插值	196
(二) 埃特金插值	197
(三) 线性拟合	200
(四) 非线性拟合	204
二、数值积分	211
(一) 变步长辛普森方法求积	211
(二) 龙贝格方法求积	212
(三) 高斯方法求积	215
(四) 切比雪夫方法求积	216
三、常微分方程的数值解法	218
(一) 定步长龙格-库塔方法	218
(二) 变步长龙格-库塔方法	221
(三) 定步长基尔方法	224
(四) 定步长哈明方法	227
(五) 病态方程组的数值解法	231
四、线性代数计算	235
(一) 列主元高斯消去法	235
(二) 高斯-约当消去法	237
五、多目标线性规划问题	240
主要参考书目	246

第一章 绪 论

随着微型计算机的普及和发展,越来越多的人掌握了计算机的使用技巧。计算机技术在各行业中已表现出无穷的魅力,生命科学的研究和应用领域早就运用了这一技术,只不过以前因为计算机技术的复杂和不普及,只能为拥有专门知识和设备的少数人使用。现在只要掌握简单的编程技巧,便可以在微型计算机上执行生物学计算、模拟、数据库操作、图像处理等技术。如果掌握了一些基本的数学和电子学基础,便可以将计算机在生物学研究中的应用扩展到更广阔的领域。

未来的生物学家将用计算机作数据分析,其模拟能力也将大于使用传统的显微镜的能力。生物学也可能成为数学应用广泛的科学,因为生物系统包含了化学、物理过程的复杂的相互作用,而所有这些过程都能用数学描述。生命系统比人能想象出的任何系统复杂得多,其大部分都不能用物理学和化学的经典数学方法来进行分析。只有在最近,由于数字计算机的出现,生物学研究者才能用数值方法去处理这种复杂的多元系统。同时,微型计算机的简单和可操作性,又进一步刺激了对生物数学的兴趣。生物学研究者可以容易地把数学分析工具引入实验室。很多知识,如数学方法和计算机应用,已为多数生物学研究者所使用,因此,这些知识也理应成为大学生物学教学的内容。

目前国内多数生物学实验室尚未配置专门的计算机辅助设备,为使本书更具实用性和可读性,主要介绍生物学模型的模拟知识和其他不依赖外围设备,而靠简单的编程技巧就能完成的工作。目的在于激发多数生物学研究者克服对复杂数学、电子和计算机知识的恐惧感,使之能迅速进入这一领域,并同科研及工作实际联系起来。

一、概念模型与数学模型在生命科学研究中的重要性

模型及其他模拟设备在科学思维过程中总是起着重要作用,人们常常不自觉地运用模拟技术,但并未意识到模型与实际系统之间的重要差别。模型可以定义为实际系统的任一表征,它既可以代表系统的结构,也可以代表系统的功能。它可以用文字、图形、数学符号或物理结构来表示系统。模型常常是概念、假说或模拟的同义词。因为模型不可能完全表示实际系统的每一个细节,所以总是含有不同程度的简化。很明显,所有科学,其本质都是探讨有关客观世界的概念模型的形成、检验和改进。重温某些代表生命系统的概念模型将是会有价值的。

作为元素结构单位的原子及作为化学反应单位的分子,在很长一段时间内仅作为概念模型存在。只是近几十年来,才能用电子显微镜直接观察一些很大的分子。X射线衍射技术,使人们能根据分子的X射线衍射图建立分子的衍射模型。X射线衍射技术仍然是一种间接的观察技术,绝不能进行直接观察。现在已经发现, α -螺旋结构能够在某种



程度上代表大多数蛋白质多肽链,但 L. Pauling (1977) 研究 α -螺旋结构时,他用纸模型做了各种各样的分子构型,直到达到了使其纸模型的重复距离与 X 射线衍射数据一致时为止。另一个典型的例子是脱氧核糖核酸 (DNA) 的双螺旋结构。这种螺旋结构是由 Waston 和 Crick 提出来的,它是建立在其分子模型已与很多的已知 DNA 的性质相符合的基础上的 (Waston, 1968)。DNA 模型激发了人们对模型和模拟的研究兴趣,至今不衰。

通常所说的典型的动物细胞或植物细胞,实际上是关于细胞的模式图。这种模式图是建立在对各种各样的细胞、用各种各样的观测方法、在多次观测的基础上的 (Hardm, 1966)。体内大多数器官的三维结构是用体视技术对千百张二维照片进行一系列分析而得到的 (Eliasand Pauly, 1966)。所得到的三维器官模型对了解在正常和异常条件下器官的功能是非常重要的。

基因的概念模型经历了长期的演变过程 (Glass, 1963)。“一根线上的一串珠子”的遗传微粒概念取代了遗传的融合理论。因此基因被看作单一的多肽编码的染色体 DNA 的一部分,尽管在这个领域内取得了很大进展,但是占支配地位的仍然是对遗传效应进行大量观察而建立起来的概念模型。每一种新模型都引导出一系列新问题,这些问题又导致对基因性质的更深入的了解。

在生态学中,食物链及食物数量金字塔是重要的概念模型,它已用来解释生态系统中能量流及物质循环。进化论是建立在大量证据的基础上的,但仍需将它视为一个概念模型。

生物化学中有各种各样有关酶作用的概念模型,包括附着于底物、酶对温度和 pH 变化的响应、糖原酵解、Krebs 循环以及光合作用中的碳循环等代谢途径,实际上都是一些概念模型。它们与其他模型一样都需要进行不断的修正,使之与新的发现相一致。每一种模型都会因其提出一些新的问题而导致实验活动的突然高涨。我们根据这些模型作出医疗、药物作用、毒剂处理及动植物营养等方面的决策。

很明显,生物学的许多成果得益于把实际的系统转化成各种各样的模型。最早期的一些模型,是对所研究系统的形态进行仔细观察及详细绘图的。这种方法现在仍然是一种对结构细节进行仔细观察的最有价值的方法。精确绘出某些细微的生物结构图就会发现,在这之前,人们尚没有认真地研究过被研究的对象。往往对这一问题产生更深刻的认识或新的了解。这说明,模型的巨大优点,常常直接来自于模型建立阶段的思维过程。很显然,建立模型对了解大多数生物系统起了重要的作用。

模型本身一般是不严密的,它可能不太明确,不同的人可能有不同的解释。为了克服这种弊病,已经研究出一种方法将概念模型转化成更容易精确描述、求值及证实的形式,这种形式的概念模型叫数学模型。

数学模型,从广义上讲,是人们考虑某个过程中某些变量间关系的总称。这种关系,不管是动态的或静态的,通常都可以用公式、表格或图形来表述。人们把这些公式、表格或图形称为数学模型。由于电子计算机的广泛应用,图形和表格均可以回归成数学解析表达式 (初等函数、非初等函数、矩阵等)。从狭义上讲,数学模型专指那些数学表达式。当然,用纯解析方法处理那些复杂的生物学问题存在很多困难,但随着应用数学及计算机的发展,计算方法逐渐丰富,计算手段逐渐增多,生命科学中数学模型



的建立将逐步完善。

建立生命科学的数学模型是非常必要的,然而大多数生物学问题无法在实验室建立其物理模型。计算机可以模拟很多动态系统的运行情况,而计算机模拟的第一步就是要建立相应的数学模型。数学模型的建立依赖于这样一个事实:假定一个过程可以用有限的变量来描述。在任一具体时刻内,这些变量的值形成了对系统描述的矢量,即静态表示。过程的动态表示则需要对所有变量之间的联系进行描述。任何一个系统的模型化都需要去发现与这些变量相联系的数学公式或函数表达式。

数学模型或模型化过程的几种职能:它概括了人们感兴趣的事物过程并常常增进人们对事物的了解;它可能提出以前被忽视的一些基本方法,也可能把人们的注意力引向临界点、临界区域或临界关系。一旦确定下数学方法,并确信这些数学公式表达了事物的主要方面,人们就可以着手解决具体问题。例如,可以在模型上解决预测、规划、插值、归纳、优化等问题。

二、计算机模拟与模型的建立步骤

最简单的模拟是在计算机上执行数学模型以产生模拟数据。用这种方法把数学模型的输出与从实际系统得到的试验数据进行比较,以便对模型作出评价。因为多元系统的模型涉及复杂的相互关系,所以模拟对了解多元系统模型特别重要。

建立模拟模型的过程迫使研究者简单地描述系统。用这种方法建立模型时,研究者须仔细考虑一般不太注意的系统细节。建立任何一种模型的目标都在于提供一种方法以获得对系统规律的新见解。模拟在这一点上是有帮助的,因为它允许将模型与试验数据进行对照,以产生可证实的结论。

用数学模型进行模拟有两种基本方法:一种方法是把数学方程简化成类似的电子器件,如电阻、电容及放大器,并用计算机来解数学方程。计算机输出的是随时间变化的电压,它即可以用 X-Y 记录仪记录,也可以用示波器显示。在研究描述如生长过程或能量这样的复杂微分方程的积分时,这种方法对模拟模型特别有用。然而,并非所有的系统都适合于这种模拟方法,而且编制程序需要有许多电子学知识。另一种方法是由数字计算机用数值方法对系统进行模拟。数字计算机模拟的方法通常包括以下步骤:

第一步须对系统进行分析,以确定建立概念模型所需要的模块。通常这种分析可以得出系统的方块图。

第二步是定义关键变量,并将每个关键变量表示为系统其他变量的函数式,然后导出方程,建立起诸变量间的实际数学关系。这种推导既可凭经验,即通过使用如曲线拟合这样的统计方法,也可用解析法,即由理论考虑而导出方程。

第三步是编制数学表达式的计算机程序并输入计算机,给各种速率常数或系数赋值,然后运行模拟程序并产生出一组模拟数据,再将模拟数据与试验数据进行比较以证实模型的可靠性。这一步通常要对模型进行某些修正,使之更精确地模拟实际系统。当证实模型有充分的有效性及合理性后,该模型就可用来进行试验,就像在真实系统上进行试验一样。

计算机模拟只能在一定程度上逼近真实系统,逼近的程度取决于所用模型的复杂



性。计算机模拟技术可以用来模拟因时空太大而不可能进行试验研究的系统。例如，对实际的捕食动物-猎物系统的研究，要计算 10~50 年时间及 10~1000 平方公里范围内的群体密度。即使在小型封闭系统内进行群体试验，也需要花许多时间进行深入的研究才能得到有意义的结果，而在计算机上进行这种小型封闭系统的模拟，只需几秒钟就能完成。

人们从事研究的目的是得出最大限度的、真实反映实际系统的概念模型。概念模型是实际系统的写照，它建立在用取得的经验能加以解释的有效信息的基础上。在研究中常常有这样的情况，即新的生物实体是由于研究其他一些系统而发现的。即使是最粗略的观察也可以得到初步的概念模型。最初的模型可以采取模拟某些已知系统的形式。概念模型常常提出一个或一些试验，完成这些试验就可以得出该模型究竟是可以接受的还是应该否决的结论。通过试验可以得到一批数据，这些数据有助于对原始观察进行新的说明或解释。不管如何，该过程一般可改善模型。这个顺序是研究工作应经历的典型循环过程。

数学模型产生于用定量的方法将概念模型进行公式化，通常表示为描述系统对时间、温度等一些变量的响应的方程式。数学模型也可由对试验数据进行统计分析而推得。在任何情况下，将模型公式化时所用到的思想一般都会改善概念模型。这种改善了的概念模型反过来又会提出进一步的试验。反复进行定量研究，可以进一步改善数学模型和概念模型。模拟能使研究者利用强有力的研究循环。模拟的主要优点在于能对数学模型和概念模型形成中所涉及的概念提供正确的检验，所以，模拟对了解定量生物系统可起到关键的作用。此外没有任何一种技术可以对发展和检验模型提供如此强有力的帮助。模拟还能让研究者利用系统模型在正常试验范围以外进行试验。也就是说，这种方法能提出这样的问题：“如果……，系统将发生什么情况”。

计算机模拟对多重非线性相互作用系统特别有价值。这在生态及生理系统这样的多元系统的处理或控制领域中特别有用。这两种系统都含有多个变量，它们之间的相互关系包含了正、负反馈循环。在很多情况下，生态系统不能用实际系统进行试验，因为它常涉及到人类和其他生物等大规模的群体，在这样情况下，计算机模拟及系统分析法，为必须做出的管理决策提供合理的依据。

人们所关心的是怎样才能建立起复杂系统的模型，或已经建立又如何来检验它们。这个问题并不像想象的那样困难。多元系统模型由一些子模型或一些模块组成，其间有很多连锁关系。每个子模型都可以独立建立，并且可以用其模拟输出数据与真实系统的对应数据进行比较而验证。只有在证实子模型的有效性之后，它才能并入一个能对整个系统起作用的模型中去。这种方法能把复杂的问题分成若干个小问题分别加以处理。

所有系统都可以看成是由分层结构组成的。在分层结构中，系统的每一单元都含有子元，子元又含有子子元。为了建立系统的模型，不必了解系统的全部子系统是如何运行的。通常，大量的子元都被归并成一个或几个子模型，而子模型大体上确定这些子元对系统的净效应。这种归并过程简化了模型的解，而且能把主要精力集中于最重要的控制因子上。

在建立分层链的中间层次的模型时，通常把更复杂的系统列到感兴趣的系统之外。



一般把外部变量的影响近似地看成环境强迫函数而包括入模型中。如果这种近似的因素在确定模型性质时起主要作用，则将模型进行扩展，以便把这样一些层次包括进去。

认为模拟数据与用实际系统进行试验时所获得的数据须严格一致是错误的。按其定义，模拟模型并非实际系统，而且没有理由要求模拟数据与实际系统的数据逐点一致。模型能呈现出与实际系统显示的主要特征相对应的关系就可以了。还应了解，模拟不应复杂到以至于对它所研究的系统起作用甚少或根本不起作用的程度。一般，建立模型是为了帮助识别系统的主要控制因素，而过分复杂的模拟反而可能掩盖了这些因素。

建立生物学模拟的方法仍处于非常幼稚的时期。许多人或许期望从模型中取得很多有意义的结果，但都因建立模型的难度太大而感到失望。也有一些人过分强调模型的预测能力，而没有认识到模型的主要价值在于能改善概念模型。由于所讨论的系统的复杂性，生物学模型不可能有物理和化学模型那样的预测能力。

生物学家在统计分析和建立模型两个方面均受到训练，因为建立模型在涉及生物学的定量研究中会起到重要作用。生物学研究工作者为了使模型作为开发和检验概念的工具，应了解建立模型的基本方法。

三、微型计算机模型的建立

计算机最近的发展，为研究者提供了一种可用于开发和执行各种复杂模拟的工具。现在大多数学者都会运用 BASIC 或 FORTRAN 语言，了解了一些简单的原理之后，就能编出一些程序来。虽然矩阵运算不能方便地在微型计算机上执行，但它对某些类型的模型是非常有用的，一旦了解了建立模型的原理，需要时就容易适应专门的模拟语言。目前我国多数高校采用 FORTRAN 语言进行教学，为了适合多数读者，本书中的程序都采用 FORTRAN 语言进行编写。考虑到本书只是一本入门性的教材，书中尽量采用简单的数学方法。附录提供了一些常用的功能较强的数学方法的 FORTRAN 程序，有兴趣的读者可根据这些数学方法将书中的有些模型进行改写，使之更简便地采用 FORTRAN 语言表述。

第二章 建立数学模型和处理模型的基本方法

生物学研究由定性描述到定量计算,标志着生物学发展已进入一个全新的阶段。对于多数生物系统来说,一个量(X)的变化必然影响到另一个或某些量(Y)的改变(可把它看成是数学上的自变量和因变量的关系)。根据自变量和因变量之间的变化特征,这两者之间的函数关系表达式为 $Y = F(X)$ 。这样一个方程就成了生物学系统的最基本的模型。

确定 X 与 Y 之间函数关系的方法主要有两种:一种叫做解析法,它产生理论模型方程;另一种方法叫做经验法,它产生经验模型方程。经验模型的产生依赖于足够多的实验数据。用统计学的方法将几种可能的方程类型分别与实验数据相拟合,选择一个最佳拟合方程作为经验模型方程。这种模型的特点是抛开被研究系统的复杂理论,只利用实验数据就能求出模型方程。而解析法则必须依赖于对所研究对象的理论推导,得到一个解析方程,如果由试验得到的数据和由解析方程得到的数据高度相关时,就可以把由理论推导的解析方程作为被研究系统的模型。经验法和解析法得到的模型方程都须回到实际系统中去进行验证,以证实模型的有效性。

一、微分方程模型

通过因变量对自变量的变化率的研究,用解析方法把简单的微分方程变成数学模型。这是生物学系统中最简单最常见的模型建立方法。

在微生物学中有一个普遍性的假定:细菌数目越多,它们就生长越快。按照这个假定可以写出生长的简单函数方程:

$$G = F(N)$$

式中, G 是生长率, N 是群体密度。如果决定生长率是群体密度的直接函数,即生长率正比于群体密度而不涉及其他因素,则生长率方程可写成如下形式:

$$G = kN$$

式中, k 是比例常数。该方程是一条斜率为 k 的直线。这种方程的价值有限,因为人们主要关心生长过程中任意时刻的群体密度,而不关心其生长率。为了使生长率方程变成能够给出任意时刻的群体密度的形式,须把生长率方程写成微分方程形式:

$$dN/dt = kN$$

这样可以通过积分重新建立群体生长的方程。下面是建立模型过程中常用的积分公式:



$$\int a dx = a \int dx = ax + c$$

$$\int x^n dx = \frac{1}{n+1} x^{n+1} + c$$

$$\int e^x dx = e^x + c$$

$$\int e^{ax} dx = \frac{e^{ax}}{a} + c$$

$$\int \frac{dx}{a+bx} = \frac{\ln(a+bx)}{b} + c$$

因为这些积分都是不定积分, 所以积分结果都含有常数 c 。

积分过程的第一步是将含有因变量 N 的所有项集中到等号的左边, 将自变量及常数集中到等号的右边:

$$\int_{N_0}^{N_t} \frac{dN}{N} = \int_0^t k dt$$

当在两积分限间积分时, 取上限的积分减去下限的积分, 结果如下:

$$\ln N_t - \ln N_0 = k \cdot t - k \cdot 0$$

$$\text{即} \quad \ln \left(\frac{N_t}{N_0} \right) = k \cdot t$$

两边取反对数并变换, 得

$$N_t = N_0 \cdot e^{k \cdot t}$$

这就得到一个有用的模型方程。该方程表示群体初始密度为 N_0 , 增长率常数为 k 的任意时刻 t 的群体密度。该方程对应的曲线是典型的指数增长曲线。该方程在一定范围内可用来描述多种正反馈或自生系统, 包括细菌的增长、人口的增长、兔群的增长等。利用这个模型方程模拟出的典型指数增长曲线 (图 2.1) (源程序略)。

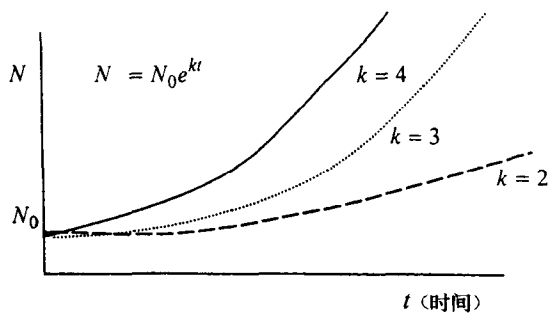


图 2.1 指数生长曲线

有些系统是随时间衰减的, 其衰减速率与任意时刻的现有量成正比。以浓度为例, 这种衰减可用下列微分方程表示:

$$\frac{dC}{dt} = -kC$$

式中, C 是某种物质浓度, 其消耗率与浓度有关, t 是时间, k 是速率常数, 负号