

MATHEMATICAL MODELING

数 模

S H U M O

SHUMO.COM 出品

2004. 1

Vol. 1 No. 1

声 明

本资料仅供参考，文章作者保留全部权力，任何将该资料全部或部分内容用于商业的行为，均须事先征得作者本人同意。转载或传播时，请保持该资料的完整性，并保留该声明。

《数 模》

MATHEMATICAL MODELING

主 办：中国数学建模网
编 辑：《数模》杂志社
主 页：www.shumo.com
电子邮件：mmjournal@yeah.net
通信地址：湖南长沙国防科大理学院
邮政编码：410073

SPONSORED BY: SHUMO.COM
COMPILED BY: *Mathematical Modeling* Editors Group
HOMEPAGE: www.shumo.com
EMAIL: mmjournal@yeah.net
ADDRESS: Science College, NUDT, Changsha Hunan
POSTALCODE: 410073

数学建模竞赛中应当掌握的十类算法

董乘宇*

(北京邮电大学, 北京 100876)

1 十类常用算法

数学建模竞赛中应当掌握的十类算法:

1. 蒙特卡罗算法。该算法又称随机性模拟算法, 是通过计算机仿真来解决问题的算法, 同时可以通过模拟来检验自己模型的正确性, 几乎是比赛时必用的方法。
2. 数据拟合、参数估计、插值等数据处理算法。比赛中通常会遇到大量的数据需要处理, 而处理数据的关键就在于这些算法, 通常使用 MATLAB 作为工具。
3. 线性规划、整数规划、多元规划、二次规划等规划类算法。建模竞赛大多数问题属于最优化问题, 很多时候这些问题可以用数学规划算法来描述, 通常使用 Lindo、Lingo 软件求解。
4. 图论算法。这类算法可以分为很多种, 包括最短路、网络流、二分图等算法, 涉及到图论的问题可以用这些方法解决, 需要认真准备。
5. 动态规划、回溯搜索、分治算法、分支定界等计算机算法。这些算法是算法设计中比较常用的方法, 竞赛中很多场合会用到。
6. 最优化理论的三大非经典算法: 模拟退火算法、神经网络算法、遗传算法。这些问题是用来解决一些较困难的最优化问题的, 对于有些问题非常有帮助, 但是算法的实现比较困难, 需慎重使用。
7. 网格算法和穷举法。两者都是暴力搜索最优点的算法, 在很多竞赛题中有应用, 当重点讨论模型本身而轻视算法的时候, 可以使用这种暴力方案, 最好使用一些高级语言作为编程工具。
8. 一些连续数据离散化方法。很多问题都是实际来的, 数据可以是连续的, 而计算机只能处理离散的数据, 因此将其离散化后进行差分代替微分、求和代替积分等思想是非常重要的。
9. 数值分析算法。如果在比赛中采用高级语言进行编程的话, 那些数值分析中常用的算法比如方程组求解、矩阵运算、函数积分等算法就需要额外编写库函数进行调用。
10. 图象处理算法。赛题中有一类问题与图形有关, 即使问题与图形无关, 论文中也会需要图片来说明问题, 这些图形如何展示以及如何处理就是需要解决的问题, 通常使用 MATLAB 进行处理。

以下将结合历年的竞赛题, 对这十类算法进行详细地说明。

2 十类算法的详细说明

2.1 蒙特卡罗算法

大多数建模赛题中都离不开计算机仿真, 随机性模拟是非常常见的算法之一。

举个例子就是 97 年的 A 题, 每个零件都有自己的标定值, 也都有自己的容差等级, 而求解最优的组合方案将要面对的是一个极其复杂的公式和 108 种容差选取方案, 根本不可能去求解析解, 那如何去找到最优的方案呢? 随机性模拟搜索最优方案就是其中的一种方法, 在每个零件可行的区间中按照正

*作者简介: 董乘宇, 曾任 SHUMO.COM 论坛“编程交流”版版主, 获 2002 年全国大学生数学建模竞赛一等奖。

态分布随机的选取一个标定值和选取一个容差值作为一种方案, 然后通过蒙特卡罗算法仿真出大量的方案, 从中选取一个最佳的。另一个例子就是去年[†]的彩票第二问, 要求设计一种更好的方案, 首先方案的优劣取决于很多复杂的因素, 同样不可能刻画出一个模型进行求解, 只能靠随机仿真模拟。

2.2 数据拟合、参数估计、插值等算法

数据拟合在很多赛题中有应用, 与图形处理有关的问题很多与拟合有关系, 一个例子就是 98 年美国赛 A 题, 生物组织切片的三维插值处理, 94 年 A 题逢山开路, 山体海拔高度的插值计算, 还有吵的沸沸扬扬可能会考的“非典”问题也要用到数据拟合算法, 观察数据的走向进行处理。此类问题在 MATLAB 中有很多现成的函数可以调用, 熟悉 MATLAB, 这些方法都能游刃有余的用好。

2.3 规划类问题算法

竞赛中很多问题都和数学规划有关, 可以说不少的模型都可以归结为一组不等式作为约束条件、几个函数表达式作为目标函数的问题, 遇到这类问题, 求解就是关键了, 比如 98 年 B 题, 用很多不等式完全可以把问题刻画清楚, 因此列举出规划后用 Lindo、Lingo 等软件来进行解决比较方便, 所以还需要熟悉这两个软件。

2.4 图论问题

98 年 B 题、00 年 B 题、95 年锁具装箱等问题体现了图论问题的重要性, 这类问题算法有很多, 包括: Dijkstra、Floyd、Prim、Bellman-Ford, 最大流, 二分匹配等问题。每一个算法都应该实现一遍, 否则到比赛时再写就晚了。

2.5 计算机算法设计中的问题

计算机算法设计包括很多内容: 动态规划、回溯搜索、分治算法、分支定界。比如 92 年 B 题用分枝定界法, 97 年 B 题是典型的动态规划问题, 此外 98 年 B 题体现了分治算法。这方面问题和 ACM 程序设计竞赛中的问题类似, 推荐看一下《计算机算法设计与分析》(电子工业出版社) 等与计算机算法有关的书。

2.6 最优化理论的三大非经典算法

这十几年来最优化理论有了飞速发展, 模拟退火法、神经网络、遗传算法这三类算法发展很快。近几年的赛题越来越复杂, 很多问题没有什么很好的模型可以借鉴, 于是这三类算法很多时候可以派上用场, 比如: 97 年 A 题的模拟退火算法, 00 年 B 题的神经网络分类算法, 象 01 年 B 题这种难题也可以使用神经网络, 还有美国竞赛 89 年 A 题也和 BP 算法有关系, 当时是 86 年刚提出 BP 算法, 89 年就考了, 说明赛题可能是当今前沿科技的抽象体现。03 年 B 题伽马刀问题也是目前研究的课题, 目前算法最佳的是遗传算法。

2.7 网格算法和穷举算法

网格算法和穷举法一样, 只是网格法是连续问题的穷举。比如要求在 N 个变量情况下的最优化问题, 那么对这些变量可取的空间进行采点, 比如在 $[a, b]$ 区间内取 $M+1$ 个点, 就是 $a, a+(b-a)/M, a+2\cdot(b-a)/M, \dots, b$ 那么这样循环就需要进行 $(M+1)^N$ 次运算, 所以计算量很大。

比如 97 年 A 题、99 年 B 题都可以用网格法搜索, 这种方法最好在运算速度较快的计算机中进行, 还有要用高级语言来做, 最好不要用 MATLAB 做网格, 否则会算很久的。

[†]编者注: 指 2002 年

穷举法大家都熟悉，就不说了。

2.8 一些连续数据离散化的方法

大部分物理问题的编程解决，都和这种方法有一定的联系。物理问题是反映我们生活在一个连续的世界中，计算机只能处理离散的量，所以需要对连续量进行离散处理。这种方法应用很广，而且和上面的很多算法有关。事实上，网格算法、蒙特卡罗算法、模拟退火都用了这个思想。

2.9 数值分析算法

这类算法是针对高级语言而专门设的，如果你用的是 MATLAB、Mathematica，大可不必准备，因为象数值分析中有很多函数一般的数学软件是具备的。

2.10 图象处理算法

01 年 A 题中需要你会读 BMP 图象、美国赛 98 年 A 题需要你知道三维插值计算，03 年 B 题要求更高，不但需要编程计算还要进行处理，而数模论文中也有很多图片需要展示，因此图象处理就是关键。做好这类问题，重要的是把 MATLAB 学好，特别是图象处理的部分。