



“十三五”科学技术专著丛书

工程领域典型算法 应用的设计与实现

高洪波 著

Design and Implementation of Typical Algorithm
Application in Engineering Field



北京邮电大学出版社
www.buptpress.com



“十三五”科学技术专著丛书

工程领域典型算法应用的 设计与实现

高洪波 著



北京邮电大学出版社
[www. buptpress. com](http://www.buptpress.com)

内 容 简 介

本书主要内容源自作者常年深入研究典型算法在实际工程领域中应用的研究成果,具有较强的原创性和实际应用的针对性。其内容主要包括马氏距离判别算法应用研究、TOPSIS 算法应用研究、灰色预测算法应用研究、灰色关联算法应用研究、高斯-牛顿非线性算法应用研究、时间序列分析算法应用研究、主成分分析算法应用研究、遗传算法应用研究、分形算法应用研究、组合预测算法应用研究。

本书对工程技术人员在工程领域中解决实际问题的算法选取、设计与实现有着较大的参考价值和借鉴意义,也可作为高等院校学生参加全国大学生数学建模竞赛等各类数学建模竞赛赛前辅导用书。

图书在版编目 (CIP) 数据

工程领域典型算法应用的设计与实现 / 高洪波著. -- 北京: 北京邮电大学出版社, 2018.12
ISBN 978-7-5635-5662-5

I. ①工… II. ①高… III. ①计算机算法 IV. ①TP301.6

中国版本图书馆 CIP 数据核字 (2018) 第 284918 号

书 名: 工程领域典型算法应用的设计与实现

责 任 编 辑: 满志文

出 版 发 行: 北京邮电大学出版社

社 址: 北京市海淀区西土城路 10 号(邮编:100876)

发 行 部: 电话: 010-62282185 传真: 010-62283578

E-mail: publish@bupt.edu.cn

经 销: 各地新华书店

印 刷: 保定市中华美凯印刷有限公司

开 本: 787 mm×1 092 mm 1/16

印 张: 8.75

字 数: 214 千字

版 次: 2018 年 12 月第 1 版 2018 年 12 月第 1 次印刷

ISBN 978-7-5635-5662-5

定 价: 35.00 元

• 如有印装质量问题,请与北京邮电大学出版社发行部联系 •

前 言

算法是工程技术人员构建模型,解决工程领域中实际问题必不可少的技能。作为工程技术人员,只有掌握算法的设计与实现的基本知识和技能,才能正确地对工程领域中的实际问题进行有效的分析和解决,保证所设计的算法模型既精准可靠又经济合理。作为工程技术人员,不仅需要较高的专业技能,而且也需要具备一定的应用计算机信息技术的能力,尤其是针对实践工作中的问题,需要具备通过选择合理的算法构建相应模型的能力,从而使解决问题的方式和手段最优,费用更省,途径更简约等,这样才能有效地提升工作效率和质量。目前,国家对高素质职业人才的需求日增,鉴于此,本书作者将自己多年来在工程领域中应用的算法设计与实现的研究成果结集成册,形成了这本专著,相信对提升工程技术人员在实际工程领域中解决实际问题典型算法的选择、模型的设计与实现有着较大的借鉴价值。

本书研究探讨了工程领域中解决实际问题典型算法应用的设计与实现,在算法选择和模型设计的实现方面,以“实用”为度,以“合理”为准,理论紧密联系实际,充分体现了工程适用和相关算法理论相结合的研究内容体系,按照认知规律兼顾工程实用的原则展开研究,本书主要有以下特点:

(1)研究内容的实用性。力求贯彻理论联系实际的原则,突出相关算法理论的实际应用,加强针对性和实用性,并尽量反映国内外最新成就和发展趋势。重点研究了典型算法在实际工程领域中应用的设计与实现。

(2)研究成果的原创性、创新性和实际应用的针对性。本书作者有在教育系统和水利工程系统工作的经历,具有丰富的在实际工程领域中典型算法应用的研究经验。本书是作者多年来潜心研究工程领域算法应用成果的集中体现,书中大多数内容是由作者近年来在北大图书馆中文核心刊物上发表的论文修改而来,因此,本书内容具有较强原创性、创新性和实际应用的针对性。

总之,本书具有较为鲜明的特点,针对实际工程中的问题,在充分考虑实际应用场景和背景的基础上,本着不求高深而求实用、适用,不求全面而求有效、高效的原则,从问题场景和相关算法的选取与设计分析入手,研究了解决工程

领域中实际问题的典型算法应用的设计与实现,其中许多都是来自实际工程领域典型算法设计与实现的应用案例,因此对于工程领域中实际问题的解决有着较大的参考意义与价值。

本书的完稿得到了作者在南京邮电大学做访问学者期间的导师、中国通信学会物联网委员会秘书长、南京邮电大学物联网学院院长、博士生导师张登银研究员的悉心指导和鼓励,使本书在学术水平和质量上都得到了进一步完善和提升。另外,本书一些工程领域中解决问题算法的选取与模型的构建研究得到了杭州汽轮机股份有限公司工程师高骥同志的倾力支持,使本书中许多研究成果更加贴近工程领域应用实际和更“接地气”。作者在此一并对他们的帮助致以诚挚的感谢!

作 者

2018年6月30日

目 录

第 1 章 绪论	1
1.1 算法概述	1
1.1.1 算法溯源	1
1.1.2 算法分类	2
1.1.3 算法的特征	2
1.1.4 算法要素	3
1.1.5 算法效能评价	3
1.1.6 算法的描述方式	3
1.1.7 基本算法	4
1.1.8 软件—程序—算法之间的关系与区别	4
1.2 算法在实际工程中的应用	5
1.2.1 建筑工程	5
1.2.2 船舶建造	5
1.2.3 金融领域	5
1.2.4 资源开发	6
1.3 算法优化理论	6
1.3.1 算法优化的概念	6
1.3.2 古典优化理论	6
1.3.3 现代传统优化理论	7
1.3.4 传统优化理论的局限性	8
1.3.5 智能优化算法及其局限性	8
1.4 实际工程领域中优化问题常见的几种典型算法	10
1.4.1 马氏距离判别算法	10
1.4.2 TOPSIS 预测算法	10
1.4.3 灰色预测算法	11
1.4.4 灰色关联算法	11
1.4.5 高斯—牛顿非线性算法	11
1.4.6 时间序列分析算法	11
1.4.7 主成分分析算法	12
1.4.8 遗传算法	12

1.4.9 分形算法	13
1.4.10 组合预测算法	13
第2章 马氏距离判别算法应用研究	15
2.1 相关理论分析	15
2.1.1 常见的距离算法与常见的相似度算法	15
2.1.2 距离算法与相似度算法的对比	15
2.1.3 距离算法与相似度算法的区别	16
2.2 马氏距离及判别法	16
2.3 基于马氏距离判别法的企业资信评估研究	17
2.4 基于主成分分析法和马氏距离判别法的故障诊断研究	22
本章参考文献	26
第3章 TOPSIS 算法应用研究	27
3.1 TOPSIS 算法及相关理论分析	27
3.1.1 TOPSIS 算法概述	27
3.1.2 TOPSIS 算法的一般实施步骤	28
3.1.3 TOPSIS 算法的缺点与改进	29
3.1.4 德尔菲法(Delphi)概述	30
3.2 基于 MATLAB 和 TOPSIS 算法的企业物流招标评估方法研究	32
本章参考文献	36
第4章 灰色预测算法应用研究	37
4.1 灰色系统理论与分析	37
4.2 灰色系统 GM 模型	38
4.3 灰色预测	39
4.3.1 灰色预测的方法	40
4.3.2 灰色预测的步骤	40
4.4 基于灰色理论的 A 级汽车销量影响因素分析与预测研究	40
4.5 基于多变量灰色模型的物流需求预测研究	44
本章参考文献	48
第5章 灰色关联算法应用研究	49
5.1 关联分析	49
5.2 灰色关联分析法	50
5.3 灰色关联分析的计算步骤	51
5.4 灰色关联算法的优缺点	51
5.5 灰色关联分析的进一步改进	52

5.6 基于灰色关联算法的物联网安全状态评估模型研究	52
5.7 江苏农机总动力影响因素灰色关联分析及 GM(1,1) 预测研究	56
本章参考文献	60
第 6 章 高斯—牛顿非线性算法应用研究	61
6.1 线性与非线性回归	61
6.1.1 线性回归与非线性回归	61
6.2 高斯—牛顿非线性算法模型	62
6.3 基于高斯—牛顿法的丹江口水库泥沙预报模型研究	63
6.4 基于高斯—牛顿法的远程教育学习兴趣趋避度模型研究	66
本章参考文献	72
第 7 章 时间序列分析算法应用研究	73
7.1 关于时间序列模型的基本理论	73
7.1.1 概述	73
7.1.2 时间序列模型的建立	73
7.2 常见的时序模型分析	74
7.2.1 时间序列模型分类	74
7.2.2 常见的时序模型概述	74
7.3 基于时间序列法的汽车销售量预测研究	78
7.4 基于改进指数平滑法的车联网交通流量预测	86
本章参考文献	91
第 8 章 主成分分析算法应用研究	92
8.1 主成分分析法概述	92
8.2 基于主成分分析算法的社区居民对社区教育满意度评价模型研究	92
8.3 基于主成分分析算法的高职课堂教学质量评价研究	96
本章参考文献	100
第 9 章 遗传算法应用研究	101
9.1 遗传算法概述	101
9.1.1 遗传算法的发展历程	101
9.1.2 遗传算法的理论研究概况	102
9.1.3 遗传算法的应用领域分析	102
9.2 遗传算法的概念、特点及实现步骤	104
9.3 基于遗传算法的物流配送网点选址研究	107
本章参考文献	111

第 10 章 分形算法应用研究	112
10.1 分形理论	112
10.1.1 分形理论概述	112
10.1.2 分形几何	113
10.1.3 分形理论的原则与意义	114
10.2 基于分形插值算法的水库泥沙淤积量模型构建研究	114
10.3 铁路货运量分形插值拟合仿真与预测研究	119
本章参考文献	122
第 11 章 组合预测算法应用研究	124
11.1 概述	124
11.2 基于组合预测算法的水库泥沙淤积量预测模型构建研究	125
本章参考文献	129

第 1 章 绪 论

1.1 算法概述

1.1.1 算法溯源

算法在我国最早可追溯至《周髀算经》《九章算术》，在中国古代文献中称为“术”。例如，在《九章算术》中给出了四则运算、最大公约数、最小公倍数、开平方根、开立方根等计算方法。又如，三国时代刘徽就给出了求圆周率的算法，即刘徽割圆术。总之，中国自唐代以来，历代都有许多专门论述“算法”的专著，数量可以用不胜枚举来形容。我国古代的筹算口诀与珠算口诀及其执行规则就是算法的雏形。古希腊数学家欧几里得在公元前 3 世纪就提出了一个算法，来寻求两个正整数的最大公约数，这就是有名的欧几里得算法，也称辗转相除法。

算法的英文单词 Algorithm 来自 9 世纪波斯数学家名字 al-Khwarizmi，因为 al-Khwarizmi 在数学上提出了算法这个概念。“算法”原为 Algorism，其含义是阿拉伯数字的运算法则，在 18 世纪演变为“Algorithm”。欧几里得算法被人们认为是史上第一个算法。第一次编写程序是 Ada Byron 于 1842 年为查尔斯·巴贝奇 (Charles Babbage) 分析机编写求解伯努利方程的程序，因此 Ada Byron 被大多数人认为是世界上第一位程序员。因为巴贝奇未能完成他的巴贝奇分析机，因此这个算法也未能在巴贝奇分析机上执行。19 世纪和 20 世纪早期的数学家、逻辑学家在定义算法上出现了困难。20 世纪的英国数学家图灵提出了著名的图灵论，并提出一种假想的计算机的抽象模型，这个模型被称为图灵机。图灵机的出现解决了算法定义的难题，图灵的思想对算法的发展起到了重要的作用。这也是图灵奖为计算机科学方面的最高奖项的缘由之一。

在古代，计算通常是指数值计算，在现代，计算已经远远地突破了数值计算的范围，包括大量的非数值计算，例如检索、表格处理、判断、决策、形式逻辑演绎等。在 20 世纪以前，人们普遍地认为，所有的问题类都是有算法的。20 世纪初，数学家们发现有的问题类是不存在算法的，遂开始进行能行性研究，于是现代算法的概念逐步明确起来。随后数学家们提出了递归函数、图灵机等计算模型，并提出了丘奇—图灵论题 (见可计算性理论)，这才有可能把算法概念形式化。按照丘奇—图灵论题，任意一个算法都可以用一个图灵机来实现；反之，任意一个图灵机都表示一个算法。按照上述理解，算法是由有限多个步骤组成的，它有下列两个基本特征：每个步骤都明确地规定要执行何种操作；每个步骤都可以被人或机器在有限的时间内完成。人们对于算法还有另一种不同的理解，要求算法除了上述两个基本特征外，还要具有第三个基本特征：虽然有些步骤可能被反复执行多次，但是在执行有限多次

之后,就一定能够得到问题的解答。也就是说,一个处处停机(即对任意输入都停机)的图灵机才表示一个算法,而每个算法都可以被一个处处停机的图灵机来实现。

1.1.2 算法分类

算法是指解题方案的准确而完整的描述,是一系列解决问题的清晰指令,算法代表着用系统的方法描述解决问题的策略机制。目前国内外有关的研究和科学文献中对于算法分类这个术语还没有明确定义,算法可以简单地根据算法设计原理、算法的具体应用和其他一些特性进行分类。可分为基本算法或根据具体应用领域进行分类。在学习中,按照学习方式,常把算法分为监督学习算法、非监督学习算法及半监督学习算法;按照图论的算法进行分类,算法可以分为哈夫曼编码、树的遍历、最短路径算法、最小生成树算法、最小树形图、网络流算法、匹配算法。

算法中指令描述的是一个计算,当其运行时能从一个初始状态(也可能为空)和初始输入开始,经过一系列有限而清晰定义的状态,最终产生输出并停止于一个终态。一个状态到另一个状态的转移不一定是确定的。随机化算法包含了一些随机输入。也就是说,能够对一定规范的输入,在有限时间内获得所要求的输出。如果一个算法有缺陷,或不适合于某个问题,那么这个算法将不会解决这个问题。不同的算法可以用不同的时间、空间或效率来完成同样的任务。一个算法的优劣可以用空间复杂度与时间复杂度来衡量。综上所述,算法可以宽泛地分为以下三类。

(1) 有限的确定性算法。这类算法在有限的一段时间内终止。它们可能要花很长时间来执行指定的任务,但仍将在一定的时间内终止。这类算法得出的结果常取决于输入值。

(2) 有限的非确定算法。这类算法在有限的一段时间内终止。然而对于一个(或一些)给定的数值,算法的结果并不是唯一的或确定的。

(3) 无限的算法。是指由于没有定义终止的条件,或定义的条件无法由输入的数据满足而不终止运行的算法。通常,无限算法的产生是由于未能确定的定义终止条件。

1.1.3 算法的特征

算法一般应具有以下特征。

(1) 输入项:一个算法有零个或多个输入,以刻画运算对象的初始情况。

(2) 确定性:算法的每一个步骤必须要确切地定义,即算法中所有有待执行的动作必须严格而不含混地进行规定,不能有歧义性。例如,欧几里得算法中,首先明确规定了“以 m 除以 n ”,而不能有“以 m 除以 n ”或“以 n 除以 m ”,这类有两种可能做法的规定。

(3) 有穷性:一个算法在执行有穷步骤后必须结束。也就是说,一个算法,它所包含的计算步骤是有限的。例如,在欧几里得算法中, m 和 n 均为正整数,在执行步骤 1 之后, r 必小于 n ,若 r 不等于 0,下一次进行步骤 1 时, n 的值已经减小,而正整数的递降序列最后必然要终止。因此,无论给定 m 和 n 的原始值有多大,步骤 1 的执行次数都是有穷的。

(4) 输出项:一个算法有一个或多个输出。输出是与输入有某个特定关系的量,简单地说就是算法的最终结果。

(5) 可行性:算法中有待执行的运算和操作必须是相当基本的。换言之,它们都是能够精确地进行的算法,执行者甚至不需要掌握算法的含义,即可根据该算法的每一步骤的要求进行操作,并最终得出正确的结果。

1.1.4 算法要素

算法要素主要包含数据的运算和操作等方面。计算机可以执行的基本操作是以指令的形式描述的。一个计算机系统能执行的所有指令的集合,称为该计算机系统的指令系统。一个计算机的基本运算和操作有四类:第一,算术运算,即加减乘除等运算;第二,逻辑运算,即与、或、非等运算;第三,关系运算,即大于、小于、等于、不等于等运算;第四,数据传输,即输入、输出、赋值等运算。另外就是算法的控制结构,一个算法的功能结构不仅取决于所选用的操作,而且还与各操作之间的执行顺序有关。

1.1.5 算法效能评价

算法效能评价包含以下五个方面。

- (1) 时间复杂度。算法的时间复杂度是指算法需要消耗的时间资源。
- (2) 空间复杂度。算法的空间复杂度是指算法需要消耗的空间资源。其计算和表示方法与时间复杂度类似,一般都用复杂度的渐近性来表示。
- (3) 正确性。算法的正确性是评价一个算法优劣最重要的标准。
- (4) 可读性。算法的可读性是指一个算法可供人们阅读的程度。
- (5) 健壮性。算法的健壮性是指一个算法对不合理数据输入的反应能力和处理能力,也称为容错性。

1.1.6 算法的描述方式

1. 用自然语言描述算法

前面关于欧几里得的算法以及算法实例的描述,使用的都是自然语言。自然语言是人们日常所用的语言,如汉语、英语、德语等。使用这些语言不用专门训练,用该法描述的算法也通俗易懂。

2. 用流程图描述算法

流程图是描述算法的常用工具。流程图用一些图形符号来表示算法。

3. 用伪代码描述算法

伪代码是用介于自然语言和计算机程序设计语言之间的文字和符号来描述算法的工具。它不用图形符号,因此书写方便、格式紧凑,易于理解,便于向计算机程序设计语言过渡。

4. 用程序描述

程序是算法用某种程序设计语言的具体实现。例如,操作系统是一个在无限循环中执行的程序,操作系统的各种任务可看成是单独的问题,每一个问题由操作系统中的一个子程序通过特定算法来实现。该子程序得到输出结果后便终止。

要使计算机能完成人们预定的工作,首先必须为预定的工作设计一个软件算法,然后再根据软件算法编写程序。计算机程序要对问题的每个对象和处理规则给出正确详尽的描述。其中程序的数据结构和变量用来描述问题的对象;程序结构、函数和语句用来描述问题的算法。算法和数据结构是程序的两个重要方面。算法是问题求解过程的精确描述,一个算法由有限条可完全机械地执行的、有确定结果的指令组成。指令正确地描述了要完成的

任务和它们被执行的顺序。计算机软件算法指令所描述的顺序执行算法的指令能在有限的步骤内终止,或终止于给出问题的解,或终止于指出问题对此输入数据无解。

1.1.7 基本算法

1. 递推法

递推法是利用问题本身所具有的一种递推关系求问题解的一种方法。它把问题分成若干步,找出相邻几步的关系,从而达到目的,此方法称为递推法。

2. 递归法

递归指的是一个过程:函数不断引用自身,直到引用的对象已知。

3. 穷举搜索法

穷举搜索法是对可能是解的众多候选解按某种顺序进行逐一枚举和检验,并从中找出符合要求的候选解作为问题的解。

4. 贪婪法

贪婪法是一种不追求最优解,只希望得到较为满意解的方法。贪婪法一般可以快速得到满意的解,因为它省去了为找最优解要穷尽所有可能而必须耗费的大量时间。贪婪法常以当前情况为基础作最优选择,而不考虑各种可能的整体情况,所以贪婪法不需要回溯。

5. 分治法

分治法是把一个复杂的问题分成两个或更多的相同或相似的子问题,再把子问题分成更小的子问题,直到子问题可以简单地直接求解,原问题的解即子问题的解的合并。

6. 动态规划法

动态规划是一种在数学和计算机科学中使用的、用于求解包含重叠子问题的最优化问题的方法。其基本思想是,将原问题分解为相似的子问题,在求解的过程中通过子问题的解求出原问题的解。动态规划的思想是多种算法的基础,被广泛应用于计算机科学和工程领域。

7. 迭代法

迭代法是数值分析中通过从一个初始估计出发寻找一系列近似解来解决问题的过程。为实现这一过程所使用的方法统称为迭代法。

8. 分支界限法

与贪婪法一样,这种方法也是用来为组合优化问题设计求解算法的,所不同的是它在问题的整个可能解空间搜索,所设计出来的算法虽然时间复杂度比贪婪法高,但是它的优点是与穷举法类似,都能保证求出问题的最佳解,而且这种方法不是盲目的穷举搜索,而是在搜索过程中通过限界,可以中途停止对某些不可能得到最优解的子空间进一步搜索(类似于人工智能中的剪枝),故它比穷举法效率更高。

1.1.8 软件—程序—算法之间的关系与区别

通常认为,软件=程序+文档=数据结构+算法+文档。另外,软件是包含程序的有机集合体,程序是软件的必要元素。任何软件都有至少一个可运行的程序。比如:操作系统给的工具软件计算器等,很多都只有一个可运行程序。而 Office 是一个办公软件包,包含了许多可运行程序的组件,如电子表格、文字处理、PPT 等。严格来说,程序是指用编程语言

编制的完成特定功能的软件。程序从属于软件,软件除包含程序外,还包含各种资料文档等。

软件是程序以及开发、使用和维护所需要的所有文档的总称,而程序是软件的一部分。算法是程序的灵魂,一个需要实现特定功能的程序,实现它的算法可以有很多种,算法的优劣决定着程序的好坏。程序员利用熟练掌握了的程序设计语言的语法,进行程序设计和软件开发就是设计好的算法。好的算法加上软件工程的理论才能做出较好的系统。软件是包含程序的有机集合体,程序是软件的必要元素,是软件的一部分。一般地,一款软件具有各种各样的功能,而程序只执行专一的命令。软件一般是由很多程序组成的,每条程序在其中做着比较固定的工作。

综上所述,程序是软件的内在因子,而软件是一个或多个程序通过编译器编译出来的成品。软件是许多能实现某些固定任务的程序的集合。也就是说,软件是由许许多多的程序组合而成的。程序是由编程人员通过某种编程语言编写出来能实现某些固定任务的代码。编程人员能够通过某种程序设计语言编写出一些能实现某些固定任务的程序段,再把这些程序段集合起来,通过编译程序编译成软件,也就是我们通常在计算机上使用的各种软件了。

1.2 算法在实际工程中的应用

随着计算机技术突飞猛进的发展,算法在解决实际工程问题中的应用越来越广泛,如在社会经济计划、工程设计、生产管理、交通运输、国防等国民经济重要领域都发挥着十分重要的作用。下面主要介绍算法在工程中具有代表性的应用。

1.2.1 建筑工程

软件算法目前已经很好地运用于工程建筑领域。许多建筑工程单位利用计算机的软件算法进行相关的成本预算、收益预算以及采购预算等。相关的建筑单位可以根据特定的程序,对所采用的数据进行输入,在完成输入后,利用统一的程序计算出建筑工程中的相关数据。目前,随着计算机软件算法水平的提高,建筑工程领域对软件算法的大量运用,很大程度上提高了工程建筑的运作效率。

1.2.2 船舶建造

软件算法在船舶建造领域有着广泛的运用。在船舶建造过程中,往往通过软件算法进行合理的计算,算出所要使用的材料量。利用软件算法中的贪婪算法,可以最大限度上节省所要使用的建造材料以及资源,减少在船舶建造过程中的资源浪费。因此可以说,软件算法的广泛运用,在很大程度上解决了船舶建造过程中有关资源浪费的一系列问题。因此,在我国船舶建造过程中一般都会选择软件算法的运用。

1.2.3 金融领域

在金融领域方面利用软件算法,是近些年来的一种发展趋势。通过软件算法,可以实时地分析出现阶段金融时态的变化过程,以及掌握相关的金融数据。因此软件算法在金融领

域的运用逐步深化。现阶段,我国银行业发行的金融 IC 卡全部采用国外芯片和国际通用标准算法(金融社保卡除外),这是软件算法的一种重要的运算形式,这种方式方法的运用,无疑为我国金融银行领域提供了便利的条件、建立了便利的基础。

1.2.4 资源开发

软件算法也广泛地运用于资源开发领域过程中,资源高效率的合理开发和利用是近些年来所追求的目标,因此,利用软件算法对开采度等数据进行计算,可以很好地把握资源的开采程度,防止资源开采过度造成资源的枯竭,或者资源的开采力度不够,不能实现较大的经济效益。因此可以说,计算机软件算法在资源开采方面也有很大的作用。

软件算法不仅运用在以上所列举的四个领域,还在医学、道路设计、数学研究等多个领域有所利用和发展。近些年来,越来越多的软件算法被开发,不同的领域运用不同的软件算法,带来了极大的便利性。

1.3 算法优化理论

最优化是应用数学的一个分支,主要研究以下形式的问题:给定一个函数,寻找一个元素满足 A ,取得最小化或者最大化。这类定式有时还被称为“数学规划”或“线性规划”。许多现实和理论问题都可以建模成这样的一般性框架。最优化是一门应用相当广泛的学科,它讨论决策问题中最佳选择的特性,构造寻求最佳解的计算方法,研究这些计算方法的理论性质及实际计算表现。伴随着计算机的高速发展和优化计算方法的进步,能够解决的优化问题的规模越来越大。最优化问题广泛见于经济计划、工程设计、生产管理、交通运输、国防等重要领域,它已受到政府部门、科研机构 and 产业部门的高度重视。

1.3.1 算法优化的概念

算法优化是对算法的有关性能进行优化,如时间复杂度、空间复杂度、正确性、健壮性。在大数据时代,算法要处理的数据的数量级也越来越大,处理问题的场景也千变万化。为了增强算法的处理问题的能力,对算法进行优化是必不可少的。算法优化一般是对算法结构和收敛性进行优化。优化问题广泛地存在于生活、生产以及科研活动中。解决优化问题有传统优化方法以及智能优化算法两大类。传统优化方法有着坚实的数学理论基础,但只能解决小部分具有特殊数学特征的优化问题。智能优化算法避开传统优化方法的强数学特征,具有很强的适应性,但缺乏统一的理论框架。下面简单回顾下优化理论发展的历史。

1.3.2 古典优化理论

在现实生活中,人们总想要达到目标最优化。例如,在生产布局方面,如何选择工厂地址、合理布置机器设备以及制定有效的运输路线,以使得运输费用最少。在生产过程中,如何制定合理的生产计划、才能有效地调度设备和人力,从而使用最少的资源生产出最多的产品。在金融投资领域,如何合理地分配投资项目,从而使投资的收益最大而风险又最小。在科学研究领域,如何合理配置参数,从而使在获得充分的数据条件下实验次数最少等等。这

些问题都可以归结为优化问题。

一般地,优化问题可定义为:给定一组决策变量,存在该组决策变量的目标函数,欲求出确定的决策变量,使其为最优。决策变量的取值范围称为可行解域。目标函数值最优有最大值和最小值两种情况,通常情况下说的目标函数最优是取其最小值,这是因为对于求最大值的优化问题,只需将目标函数取反即可将其转化为求最小值的优化问题。

由于优化问题广泛存在,因此,优化问题很早就得到了人们的关注。早期的数学家对优化问题的数学特征进行了研究,根据一元函数局部优的导数特点提出了局部优的一阶必要条件,欧拉将这个条件扩展到多元函数。无约束优化问题的一阶必要条件和二阶充分条件将其转化为解方程或方程组问题。随后为了对具有等式约束条件的多元函数优化提供一般性的方法,又探索出了乘子法,乘子法是将等式约束优化问题转化为解方程组问题的优化方法。

乘子法用解方程的方法来解优化问题,使数学方法可以得到充分的应用。但是,由优化问题转化而来的方程组,往往比较复杂难解,有时候还是不定方程组。因此,用迭代算法找近似解的方法成为优化算法的主要方法。使用迭代算法找近似解,就是从单个初始解出发,向确定的方向移动一定的步长距离,如此迭代直至解的误差在设定范围内,最后的近似解就是优化问题的解。至此,古典优化理论发展到了一个“瓶颈”,直到 20 世纪中叶优化理论也没有新的进展。

1.3.3 现代传统优化理论

早期的研究包含了优化理论的两大板块——优化问题和优化算法。优化问题的理论主要集中在目标函数在可行解域上的数学特征的研究,所回答的问题是什么特征的解是优化解,而优化算法的理论则是研究如何找到具有优化特征的解,回答的问题是如何得到具有优化特征的解。由于优化问题的广泛存在,这就要求发展出解决这些问题的一般方法,要求优化理论进一步完善。数学的发展为优化理论的发展奠定了基础。传统的优化理论仍然是先研究优化问题,得到优化解的数学特征,再根据这个数学特征设计优化算法。

现代传统优化理论是从线性规划理论的提出开始大发展的,这得益于现代经济发展的需要和计算机的出现。现代经济的发展要求优化理论解决在经济社会中面临的各种优化问题,这些问题往往规模很大,靠人工计算是不可能的。只有在电子计算机出现后,这些优化理论和计算任务才成为可能。

受线性规划发展的启发,人们很自然地将目标投向非线性领域。20 世纪中叶,有关学者提出了约束优化问题的必要条件和充分条件,标志着非线性规划正式诞生。大多数非线性规划问题都是比较难以解决的,而其中的凸规划问题是相对容易的,这方面的研究大多是借鉴线性规划结果,再将其推广到凸规划领域。例如,在二次规划中的单纯形算法以及在凸规划中引入的内点法。随后,学术界又提出了动态规划原理,为优化问题的计算提供了可行的方法。这些优化理论奠定了现代优化理论与算法的基础。但在这之后,优化理论的进展主要表现在以下几个方面。

(1) 将在传统微积分领域内的结果推广到现代分析领域,形成更一般的形式优化算法;

(2) 稳定理论复杂性方面的算法研究以及对某些特定场合的优化技术研究,如平滑技术、松弛技术以及一些变换函数的方法。

较长时间以来,这些进展只是发生在局部领域,说明传统的优化理论与技术已经处于“瓶颈”的状态,需要寻找新的优化理论与技术来满足日益复杂的需求。

1.3.4 传统优化理论的局限性

传统优化理论有坚实的数学理论基础,在社会经济、科研和生产领域得到大量的应用。由于需要预先获知优化问题的数学特征,才能根据优化问题的数学特征设计出优化算法,所以传统优化理论有着以下的局限性。

首先,传统优化理论只能解决优化问题中的一小类问题,其数学特征能够被精确认识。传统优化算法是确定性的、精确的算法,其每一步的搜索都要有充分的依据,这需要建立在对优化问题的目标函数以及可行解域的充分认识之上。由于数学本身的局限性,它只能附加一些条件来降低问题的困难性。因此,数学对优化问题的模型有着很严格的条件,对其解空间有严格的数学特征的限制。比如线性规划需要目标函数和约束条件均是线性的凸规划,并要求对目标函数和约束条件的数学特征放宽一些,但要求是凸的,仍然是非常严格的条件。这些条件的限制使传统优化算法的应用范围大大变窄,只有一小类问题能够满足其条件。同时,现实中遇到的许多问题,其解空间的特征目前尚不能清楚地知道如组合优化问题等。所以,传统优化算法对于不能从数学上精确描述其特征的优化问题也是无能为力的。

其次,传统优化算法无法跳出局部优化解。传统优化算法从本质上讲是局部寻优方法,对于全局只有一个局部优化解的问题是合适的,但对于全局有多个局部优化解的情况,它就无法保证收敛到全局最优解。也是由于这个原因,它非常依赖于初始解的选取,一般只会收敛到初始解所在区域的局部最优解。

最后,传统优化方法要求先求得优化问题的优化问题的数学特征,然后再针对该特征进行设计优化算法。这无形中增加了其应用的难度,也对应用者提出了较高的要求,不但需要较为深厚的数学理论基础,而且需要对优化问题的相关领域知识要相当熟悉,同时还必须具备算法设计的能力。不仅如此,在实际的应用中,如果优化问题不能用传统优化问题直接解决,那么就需要将原问题进行转化,这就要求有很高的数学技巧。这也使传统优化算法在工程上的应用受到较大的限制,只有一些成熟的、已得到很好解决的优化问题才能使用传统优化算法。

1.3.5 智能优化算法及其局限性

1. 智能优化算法

传统优化理论的局限性部分地反映了数学理论的局限性。以目前的数学理论尚不能给出大多数的优化问题的优化问题的数学特征。所以在数学理论尚未有实质性突破的情况下,传统优化理论的局限性难以得到克服。近年来,优化算法研究者们并没有局限在传统优化方法的研究中,他们开辟了另一条优化计算的道路——不需要预先确知优化问题的优化问题的数学特征,通过启发式的算法对优化问题求解。启发式优化算法是近似算法,不保证求得最优解,并在有限的时间内求得可以接收的近似优化解。启发式优化算法以其快速有效的特点得到广泛应用,也因其无法得到最优解受到颇多诟病。

智能优化算法就是一类特殊的启发式优化算法,是通过模仿自然和生物现象发展出一