

国外信息技术优秀图书选译

近似算法

Approximation Algorithms

Vijay V. Vazirani 著

郭效江 方奇志 农庆琴 译

本书系统总结了到本世纪初为止近似算法领域的成果，重点关注近似算法的设计与分析，介绍了这个领域中最重要的问题以及所使用的基本方法和思想。全书分为三部分：第一部分使用不同的算法设计技巧给出了下述优化问题的组合近似算法：集合覆盖、施泰纳树和旅行商、多向割和 k -割、 k -中心、反馈顶点集、最短超字符串、背包、装箱问题、最小时间跨度排序、欧几里得旅行商等。第二部分介绍基于线性规划的近似算法。第三部分包括四个主题：在一个格中找一个最短向量、计数问题的可近似性、基于 PCP 定理的近似困难性以及未解决的问题等，这些问题都是近似算法领域中的前沿研究内容。

本书可作为计算机科学、应用数学、运筹学、信息科学与网络工程、物流与交通运输、管理科学与工程、生命科学、电子科学与技术等学科专业的研究生及高年级本科生的教学用书，对相关领域的科学研究人员也具有参考价值。

“本书涵盖了获取难解组合最优化问题和计数问题的近似解的主要理论方法。它包括简洁优雅的组合理论，有用又有趣的算法以及组合问题所固有复杂性的深入结果。讲解清晰透彻，练习选取精当，本书必将被所有数学和算法研究者所接受并喜爱。”

——加州大学伯克利分校 Richard Karp

“很高兴推荐 Vijay Vazirani 撰写的这本书，本书关注近似算法这一重要且适时的主题，写作上乘，内容全面。我确信广大读者无论是将其用作近似性的入门教材，还是作为近似算法诸多问题的参考资料，都会发现本书极为有用。”

——微软研究院 László Lovász

<http://academic.hep.com.cn>

ISBN 978-7-04-029863-5



9 787040 298635 >

学科类别：数学、计算机

定价 49.00元

国外信息技术优秀图书选译

Approximation Algorithms

近 似 算 法

Jinsi Suanfa

Vijay V. Vazirani 著

郭效江 方奇志 农庆琴 译



高等教育出版社·北京
HIGHER EDUCATION PRESS BEIJING

图字: 01 - 2009 - 5289 号

Translation from the English language edition:

Approximation Algorithms by Vijay V. Vazirani

Copyright © Springer - Verlag Berlin Heidelberg 2001

Springer is part of Springer Science + Business Media

All Rights Reserved

图书在版编目(CIP)数据

近似算法/(美)瓦齐拉尼(Vazirani, V. V.)著;郭效江,
方奇志,农庆琴译. —北京:高等教育出版社,2010.9

书名原文:Approximation Algorithms

ISBN 978 - 7 - 04 - 029863 - 5

I. ①近… II. ①瓦… ②郭… ③方… ④农…

III. ①近似计算 - 高等学校 - 教材 IV. ①O242.2

中国版本图书馆 CIP 数据核字(2010)第 144679 号

策划编辑	刘 英	责任编辑	廖肇源	封面设计	刘晓翔
责任绘图	尹 莉	版式设计	余 杨	责任校对	俞声佳
责任印制	陈伟光				

出版发行 高等教育出版社

社 址 北京市西城区德外大街 4 号

邮政编码 100120

经 销 蓝色畅想图书发行有限公司

印 刷 涿州市星河印刷有限公司

购书热线 010-58581118

咨询电话 400-810-0598

网 址 <http://www.hep.edu.cn>

<http://www.hep.com.cn>

网上订购 <http://www.landaco.com>

<http://www.landaco.com.cn>

畅想教育 <http://www.widedu.com>

开 本 787 × 1092 1/16

印 张 23.75

字 数 500 000

版 次 2010 年 9 月第 1 版

印 次 2010 年 9 月第 1 次印刷

定 价 49.00 元

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换。

版权所有 侵权必究

物料号 29863-00

序 言

虽然这看似矛盾,但近似的想法主宰着所有精确科学。

伯特兰·罗素 (1872—1970)

大部分自然的最优化问题 (包括那些在重要应用领域中产生的最优化问题) 是 **NP**-难解的。因此,在普遍相信的 $P \neq NP$ 猜想下,获取它们的精确解所耗费的时间是令人望而却步的。通过多项式时间算法,弄清楚这些问题的可近似性因此成为计算机科学和数学中一个引人注目的科学探究主题。本书介绍现今所呈现出的近似算法理论。期待着近似算法理论随着时间而变化是合理的。

全书分为三部分。第一部分利用不同的算法设计技术对许多重要问题给出组合算法。算法设计技术的不断变化可能会使第一部分显得重点不突出。不过,这是合理的——**NP**-难解问题的种类很多,不能指望少数技巧就能求解不同的 **NP**-难解问题。当然,在这部分,我们故意不严格分类算法技术以免使内容琐碎。而是改为尽可能精确地抓住每个问题的个性,并指出这些问题和求解它们的算法之间的联系。

第二部分介绍基于线性规划的算法。这些算法可分类为两个基本技术:舍入和原始对偶模式。但是再一次,能够得到的精确近似保证依赖于所用的特定线性规划松弛,并且如同证明数学定理不存在固定的方法一样,发现好的松弛也不存在固定的方法 (熟悉复杂性理论的读者将认识到这就是 $P \neq NP$ 问题背后的哲学问题)。

第三部分包括四个重要主题。第一个主题是找格的最短向量,因为几个原因,它值得单独处理 (参见第 27 章)。

第二个主题是计数 (而非优化) 问题 (计数给定实例的解) 的可近似性。几乎所有已知的 **NP**-完全问题的计数形式都是 $\#P$ -完全的。有趣的是,除少数例外,这对 **P** 中的问题同样成立。为得到后一类问题的有效近似计数算法,已建立起一套令人赞叹的理论。这些算法中的大部分都是基于马尔可夫链蒙特卡罗 (MCMC) 方法,这个主题值得用一本书来介绍,因此本书不予介绍。在第 28 章中不使用 MCMC 方法介绍两个基本计数问题的组合算法。

第三个主题围绕着最近的突破性成果, 这些成果给出许多关键问题的近似困难性, 并对将近似算法作为一种深刻理论提供新的合理性。假定主要的技术定理 (PCP 定理) 成立, 在第 29 章中概述这些结果。不幸的是, PCP 定理目前还没有简单的证明。

第四个主题由这个年轻领域中许多未解决的问题组成。所列出的问题以正时兴的问题和议题为中心, 但肯定会有所遗漏。精确算法已经被深入地研究了四十多年, 但仍在获取基本的认识。考虑到在自然计算问题中, 多项式时间可解性是例外而不是规则, 期待着近似算法理论随着时间迅速发展是合理的。

集合覆盖问题在近似算法理论和这本书中都占据了特殊位置。它对介绍第一部分和第二部分的关键概念和一些基本算法设计技术提供了一个特别简单的背景。为了对这个重要问题给出完整的讨论, 在第三部分对它给出了近似困难性结果, 尽管证明相当复杂。这个近似困难性结果基本上匹配已知的最好算法的保证——这是介绍这个相当困难的证明的另一个原因。

与米开朗琪罗部分艺术研究工作的下述类比能很好地说明我们关于算法设计和讲解的原则。米开朗琪罗的主要工作是在采石场寻找能引起注意的石块, 然后长时间凝视它们以确定它们自然地想要成为的形态。雕刻工作以最低限度的方式完成这个形态。同样, 我们愿意开始于清晰、简单规定的问题 (或许是实际想要求解的问题的简化形式)。大部分算法设计工作实际上是要认识问题与算法相关的组合结构。算法最低限度地利用这个结构。本书中算法的讲解也将遵循这个类比, 讲解着重说明问题所提供的结构并使算法最低限度地利用这个结构。

我们努力使每章保持简短和简单, 经常只介绍关键结果。而把推广和相关的结果放入练习。练习还包括由于逻辑限制而不能详细包括的其他重要结果。对一些练习提供了提示; 但是, 一道练习的困难程度和是否对它提供提示是没有关联的。

本书适用于高年级本科生和研究生的近似算法课程。全部讲授完本书的内容至少需要两个学期, 因此, 如果只准备讲授一个学期, 那么教师在选择主题时会有很大的空间。一门算法和 NP-完全性理论的大学本科课程作为大部分章节的前提条件就够了。对于完全性, 我们提供几个主题的背景信息: 附录 A 中的复杂性理论, 附录 B 中的概率论, 第 12 章中的线性规划, 第 26 章中的半定规划和第 27 章中的格。(附录 A 中把大部分空间留给了自可归约性概念, 因为其他资料中很少提到这个概念。) 本书也能用作低年级本科生和研究生算法课程的补充教材。第一部分和第二部分的最初几章适用于这个目的。在这两部分中, 章大致上是按照困难性增加的次序排列的。

预期到广泛的读者, 我们决定不在任何 Springer 丛书中出版这本书——即使是它享有声望的“黄皮”丛书^①。(不过, 我们并不反对在封面上添加一小片黄颜色!) 当前正计划着下述翻译: 由 Claire Kenyon 译成法语, 由 Takao Asano 译

^① 指的是数学教材丛书。——译者注

成日语, 由 Ion Măndoiu 译成罗马尼亚语。欢迎来自读者的校正和评论。为了这个目的, 我们设立了一个专门的 email 地址: `approx@cc.gatech.edu`。

最后, 谈谈实用效果。对于寻找高性能 (具有在最优解的 2% 或者 5% 以内的误差) 算法的实践者来说, 在最优解的因子 2 或者更坏的因子 $O(\log n)$ 以内的算法能有多好? 更进一步, 由此看来, 近似保证的改进 (比如从因子 2 到 $3/2$) 能有用到何种程度?

我们讨论一下这两个问题并指出这些论断中的一些谬误。近似保证仅仅反映算法关于大部分病态实例的性能。或许把近似保证看成促使我们更深入地研究问题的组合结构并发现利用这个结构的更强有力工具的一种度量更合适。已注意到当得到有更好保证的算法的时候, 构造紧例子的困难性明显增大。实际上, 对于一些近来的算法, 得到紧例子已经独立地成为一篇论文 (例如, 参见节 26.7)。实验已证实这些算法和其他复杂算法对于典型实例能达到想要的 2% 到 5% 量级的误差界限, 虽然它们的最坏情形误差界限要高得多。另外, 应将已被理论证明的算法看成核心算法思想, 这个思想需要很好地融入特定应用中所产生的实例。

希望本书能成为近似算法理论发展的催化剂并产生实用效果。

致谢

本书基于在印度理工学院 (德里, 1992 年春季和 1993 年春季)、佐治亚理工学院 (1997 年春季, 1999 年春季和 2000 年春季)、离散数学与理论计算机科学研究中心 (1998 年秋季) 所教授的课程。1992 年春季课程产生了关于这个主题的第一个课堂讲义。值得指出的是, 本书的多半部分是基于后来的研究结果。

众多朋友 (包括他们的家人) 的帮助促使这本书变为现实。首先, 我想感谢 Naveen Garg, Kamal Jain, Ion Măndoiu, Sridhar Rajagopalan, Huzur Saran 和 Mihalis Yannakakis——我和他们的广泛合作使得本书所介绍的许多思想得以成型。我很幸运得到 Ion Măndoiu 对许多事情的帮助和建议——他对版面布局和图形的一流认识使得这个版式得以呈现。特别感谢你, Ion!

我想对这个领域中的众多专家表示我的感谢, 感谢他们以各种形式对这项工作的慷慨帮助, 从决定内容和它的组织, 提供完成稿的反馈, 确保参考文献的正确性和完整性到设计练习和帮助列出未解决的问题。感谢 Sanjeev Arora, Alan Frieze, Naveen Garg, Michel Goemans, Mark Jerrum, Claire Kenyon, Samir Khuller, Daniele Micciancio, Yuval Rabani, Sridhar Rajagopalan, Dana Randall, Tim Roughgarden, Amin Saberi, Leonard Schulman, Amin Shokrollahi 和 Mihalis Yannakakis, 特别感谢 Kamal Jain, Éva Tardos 和 Luca Trevisan。

还有许多其他人提供了有价值的评论和讨论。特别地, 我想感谢 Sarmad

Abbasi, Cristina Bazgan, Rogerio Brito, Gruia Calinescu, Amit Chakrabarti, Mosses Charikar, Joseph Cheriyan, Vasek Chvátal, Uri Feige, Cristina Fernandes, Ashish Goel, Parikshit Gopalan, Mike Grigoriadis, Sudipto Guha, Dorit Hochbaum, Howard Karloff, Leonid Khachian, Stavros Kolliopoulos, Jan van Leeuwen, Nati Lenial, George Leuker, Vangelis Markakis, Aranyak Mehta, Rameesh Motwani, Prabhakar Raghavan, Satish Rao, Miklos Santha, Jiri Sgall, David Shmoys, Alistair Sinclair, Prasad Tetali, Pete Veinott, Ramarathnam Venkatesan, Nisheeth Vishnoi 和 David Williamson。我肯定还遗漏了几个名字——我向他们致以歉意并同样向他们表示感谢。参加我这个主题的课程并且记录讲义的众多学生起到了特殊的作用,难以一一列举他们的名字。我想向他们一同表达我的谢意。

我想感谢印度理工学院(德里)(特别感谢 Shachin Maheshwari)、佐治亚理工学院和离散数学与理论计算机科学研究中心,感谢他们提供了舒适的、能给予支持的和学术上丰富的环境。感谢 NSF 以 CCR-9627308 和 CCR-9820896 提供的基金支持。

与 Hans Wössner 在编辑事务上的合作是愉快的。他认真对待所有需要处理的事务和他对作者独特观点的敏锐捕捉让人印象特别深刻。还要感谢 Frank Holzwarth 分享他在 L^AT_EX 上的技术。

没有家人们全力的支持是很难完成这个规模的计划的。幸运的是,我的一些家人也是研究伙伴——我的妻子 Milena Mihail 和我的兄长 Umesh Vazirani。在这个计划实施的过程中,小 Michel 的诞生带来了新的欢乐和活力,虽然他使计划的结束变得更有挑战性!最重要的是,我想要感谢我父母——我父亲是几本土木工程书籍的著名作者,我母亲对印度古典音乐有着深入的理解——感谢他们坚定的支持和鼓励。这本书献给他们。

佐治亚州,亚特兰大市,2001 年 5 月

Vijay Vazirani

目 录

1	引言	1
1.1	确定 OPT 的下界	2
1.1.1	基数顶点覆盖的近似算法	2
1.1.2	能够改进近似保证吗?	3
1.2	有好刻画的问题和最小最大关系	4
1.3	练习	6
1.4	注释	8

第一部分 组合算法

2	集合覆盖	13
2.1	贪婪算法	13
2.2	分层	15
2.3	应用于最短超字符串	17
2.4	练习	20
2.5	注释	23
3	施泰纳树和旅行商	25
3.1	度量施泰纳树	25
3.1.1	基于最小生成树的算法	26
3.2	度量旅行商	27
3.2.1	简单的因子 2 算法	28
3.2.2	改进因子到 $3/2$	29
3.3	练习	31
3.4	注释	34

4	多向割和 k -割	35
4.1	多向割问题	35
4.2	最小 k -割问题	37
4.3	练习	39
4.4	注释	42
5	k -中心	43
5.1	参数修剪应用于度量 k -中心	43
5.2	加权形式	45
5.3	练习	47
5.4	注释	48
6	反馈顶点集	49
6.1	圈加权图	49
6.2	分层应用于反馈顶点集	52
6.3	练习	54
6.4	注释	55
7	最短超字符串	57
7.1	因子 4 算法	57
7.2	改进到因子 3	60
7.2.1	达到最优压缩的一半	61
7.3	练习	62
7.4	注释	62
8	背包	63
8.1	背包的伪多项式时间算法	63
8.2	背包的 FPTAS	64
8.3	强 NP-难解性和 FPTAS 的存在性	65
8.3.1	FPTAS 是最值得找的近似算法吗?	66
8.4	练习	67
8.5	注释	67
9	装箱问题	69
9.1	渐近 PTAS	70
9.2	练习	71

9.3 注释	72
10 最小时间跨度排序	73
10.1 因子 2 算法	73
10.2 最小时间跨度的 PTAS	74
10.2.1 物体大小的种类固定的装箱问题	75
10.2.2 时间跨度问题归约到受限制的装箱问题	75
10.3 练习	76
10.4 注释	77
11 欧几里得旅行商	79
11.1 算法	79
11.2 正确性证明	81
11.3 练习	83
11.4 注释	84

第二部分 基于线性规划的算法

12 线性规划对偶介绍	87
12.1 线性规划对偶定理	87
12.2 最小最大关系和线性规划对偶	91
12.3 两个基本的算法设计技术	93
12.3.1 两个技术的比较和整间隙的概念	94
12.4 练习	95
12.5 注释	99
13 用对偶拟合分析集合覆盖	101
13.1 对贪婪集合覆盖算法进行基于对偶拟合的分析	101
13.1.1 能改进这个近似保证吗?	103
13.2 集合覆盖的推广	104
13.2.1 对偶拟合应用于有约束的集合多次覆盖	105
13.3 练习	108
13.4 注释	109
14 舍入应用于集合覆盖	111
14.1 简单的舍入算法	111

14.2	随机舍入	112
14.3	顶点覆盖的半整性	113
14.4	练习	115
14.5	注释	115
15	对集合覆盖使用原始对偶模式	117
15.1	模式概述	117
15.2	对集合覆盖使用原始对偶模式	119
15.3	练习	121
15.4	注释	121
16	最大可满足性	123
16.1	处理大子句	123
16.2	使用条件期望方法来去随机化	124
16.3	使用线性规划舍入来处理小子句	125
16.4	3/4 因子算法	127
16.5	练习	129
16.6	注释	129
17	无关平行机排序	131
17.1	线性规划背景下的参数修剪	131
17.2	极点解的性质	132
17.3	算法	133
17.4	极点解的附加性质	133
17.5	练习	135
17.6	注释	135
18	树的多割和树的整数多商品流	137
18.1	问题和它们的线性规划松弛	137
18.2	基于原始对偶模式的算法	139
18.3	练习	142
18.4	注释	143
19	多向割	145
19.1	令人感兴趣的线性规划松弛	145
19.2	随机舍入算法	147

19.3	结点多向割的半整性	149
19.4	练习	152
19.5	注释	155
20	一般图的多割	157
20.1	和多商品流	157
20.2	基于线性规划舍入的算法	158
20.2.1	增长区域: 连续过程	159
20.2.2	离散过程	161
20.2.3	找相继区域	161
20.3	紧例子	163
20.4	多割的一些应用	164
20.5	练习	165
20.6	注释	167
21	最稀疏割	169
21.1	需求多商品流	169
21.2	线性规划模型	170
21.3	度量, 割填装和 ℓ_1 -可嵌入性	172
21.3.1	度量的割填装	172
21.3.2	度量的 ℓ_1 -可嵌入性	173
21.4	度量的低失真 ℓ_1 -嵌入	175
21.4.1	确保不过度缩短单独的边	175
21.4.2	确保不过度缩短边	178
21.5	基于线性规划舍入的算法	178
21.6	应用	179
21.6.1	边扩展	179
21.6.2	传导率	180
21.6.3	平衡割	180
21.6.4	最小割线性排列	181
21.7	练习	182
21.8	注释	184
22	施泰纳森林	185
22.1	线性规划松弛和对偶	185
22.2	同步原始对偶模式	186

22.3	分析	190
22.4	练习	192
22.5	注释	197
23	施泰纳网络	199
23.1	线性规划松弛和半整性	199
23.2	迭代舍入技术	202
23.3	刻画极点解	204
23.4	计数论证	206
23.5	练习	208
23.6	注释	214
24	设施定位	217
24.1	对偶的直观理解	218
24.2	松弛原始互补松弛条件	219
24.3	基于原始对偶模式的算法	219
24.4	分析	221
24.4.1	运行时间	222
24.4.2	紧例子	223
24.5	练习	223
24.6	注释	226
25	k-中位点	227
25.1	线性规划松弛和对偶	227
25.2	高级想法	228
25.3	随机舍入	230
25.3.1	去随机化	232
25.3.2	运行时间	232
25.3.3	紧例子	233
25.3.4	整间隙	233
25.4	近似算法的拉格朗日松弛技术	234
25.5	练习	234
25.6	注释	237
26	半定规划	239
26.1	严格二次规划和向量规划	239

26.2	半正定矩阵的性质	240
26.3	半定规划问题	242
26.4	随机舍入算法	243
26.5	对 MAX-2SAT 改进近似保证	246
26.6	练习	248
26.7	注释	250

第三部分 其他主题

27	最短向量	255
27.1	基、行列式和正交性亏量	256
27.2	欧几里得算法和高斯算法	258
27.3	使用格拉姆 - 施密特正交化确定 OPT 的下界	259
27.4	推广到 n 维空间	261
27.5	对偶格和它的算法应用	265
27.6	练习	268
27.7	注释	272
28	计数问题	273
28.1	计数 DNF 解	274
28.2	网络可靠性	276
28.2.1	确定接近最小割的数目的上界	276
28.2.2	分析	278
28.3	练习	279
28.4	注释	282
29	近似困难性	283
29.1	归约、间隙和困难性因子	283
29.2	PCP 定理	285
29.3	MAX-3SAT 的困难性	288
29.4	变量出现次数有限的 MAX-3SAT 的困难性	289
29.5	顶点覆盖和施泰纳树的困难性	291
29.6	团的困难性	293
29.7	集合覆盖的困难性	296
29.7.1	NP 的两个证明者一轮刻画	297
29.7.2	配件	298

29.7.3 通过平行重复减小误差概率	299
29.7.4 归约	300
29.8 练习	303
29.9 注释	305
30 未解决的问题	307
30.1 有常数因子算法的问题	307
30.2 其他最优化问题	309
30.3 计数问题	310
30.4 注释	314
 附录	
A 为算法设计者概述复杂性理论	317
A.1 证据和 NP 类	317
A.2 归约和 NP-完全性	318
A.3 NP-最优化问题和近似算法	319
A.3.1 保持近似因子的归约	320
A.4 随机复杂类	321
A.5 自可归约性	322
A.6 注释	324
 B 概率论的基本事实	 325
B.1 期望和矩	325
B.2 均值偏差	326
B.3 基本分布	327
B.4 注释	327
 参考文献	 329
 问题索引	 355
 主题索引	 359

1 引言

从允许近似到任何想要的程度到本质上根本不允许近似, **NP**-难解最优化问题呈现出丰富的可能性. 尽管有这些差异, 近似算法设计过程的背后是一些共同的原理. 本章将研究这些原理.

只有当最优化问题具有与算法相关的组合结构且这个结构能被用作“立足点”来有效地追踪最优解时, 最优化问题才是多项式时间可解的. 设计多项式时间精确算法是双管齐下的处理: 解开问题中的这个结构并寻找能够利用这个结构的算法技术.

虽然 **NP**-难解最优化问题不能为有效地寻找最优解提供立足点, 但是它们仍然能为有效地寻找接近最优的解提供立足点. 因此, 在很大程度上近似算法的设计过程与精确算法的设计过程并没有很大的不同. 它依然需要解开相关的结构并寻找利用这个结构的算法技术. 通常情况下, 组合结构会更加精细, 而算法技术常常来自于精确算法研究中发展起来的一些强有力的算法工具的推广和扩展.

另一方面, 更深入地研究设计近似算法的过程, 可以看到它有其自身的普遍原理. 我们利用下面的简单背景, 在节 1.1 中举例说明部分原理.

问题 1.1 (顶点覆盖) 给定无向图 $G = (V, E)$, 顶点集上的费用函数 $c: V \rightarrow \mathbb{Q}^+$, 找最小费用顶点覆盖, 即求集合 $V' \subseteq V$, 使得每条边至少有一个端点与 V' 关联. 每个顶点有单位费用的特殊情形称为基数顶点覆盖问题.

因为近似算法的设计需要有技巧地处理 **NP**-难解性并从这一处理中获取有效近似解, 所以回顾复杂性理论的一些关键概念对读者而言是有益的. 附录 A 和节 1.3 中的一些练习就是针对这个目的而提供的.

给出 **NP**-难解最优化问题及其近似算法 (例如, 参见练习 1.9 和 1.10) 的精确定义是重要的. 但由于这些定义非常理论化, 所以把它们放入附录 A. 下面我们给出要点以便快速开始.

NP-难解最优化问题 Π 是最小化或者最大化问题. Π 的每个有效实例 I 具有一个非空可行解集合, 给每个可行解赋一个非负有理数, 这个非负有理数称为它的目标