

信息技术和电气工程学科国际知名教材中译本系列

Convex Optimization

凸优化

Stephen Boyd Lieven Vandenberghe 著
王书宁 许鋈 黄晓霖 译

Stephen Boyd and
Lieven Vandenberghe

Convex
Optimization

清华大学出版社



信息技术和电气工程学科国际知名教材中译本系列

Convex Optimization

凸优化



Stephen Boyd Steven Vandenberghe 著

王书宁 许鋈 黄晓霖 译

清华大学出版社
北京

北京市版权局著作权合同登记号 图字:01-2009-3869

Authorized translation from the English language edition, entitled Convex Optimization, ISBN 978-0-521-83378-3 by Stephen Boyd and Lieven Vandenberghe, published by Cambridge University Press, copyright © 2004.

All Rights Reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage retrieval system, without permission from Cambridge University Press, Inc.

Simplified Chinese language edition published by **TSINGHUA UNIVERSITY PRESS** Copyright © 2012.

本书中文简体版由剑桥大学出版社授权给清华大学出版社出版发行。未经许可,不得以任何方式复制或抄袭本书的任何部分。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

凸优化/(美)鲍德(Boyd, S.)等著;王书宁等译. —北京:清华大学出版社,2013.1

信息技术和电气工程学科国际知名教材中译本系列

书名原文:Convex Optimization

ISBN 978-7-302-29756-7

I. ①凸… II. ①鲍… ②王… III. ①凸分析—教材 IV. ①O174.13

中国版本图书馆CIP数据核字(2012)第190148号

责任编辑:王一玲

封面设计:常雪影

责任校对:李建庄

责任印制:李红英

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦A座 邮 编:100084

社总机:010-62770175 邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质量反馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

课件下载: <http://www.tup.com.cn>, 010-62795954

印装者:清华大学印刷厂

经 销:全国新华书店

开 本:185mm×260mm 印 张:44.75 字 数:1115千字

版 次:2013年1月第1版 印 次:2013年1月第1次印刷

印 数:1~3000

定 价:99.00元

产品编号:031849-01

译者序

本书由美国斯坦福大学Stephen Boyd教授和加州大学洛杉矶分校Lieven Vandenberghe教授合著，从理论、应用和算法三个方面系统地介绍凸优化内容。

凸优化在数学规划领域具有非常重要的地位。从应用角度看，现有算法和常规计算能力已足以可靠地求解大规模凸优化问题，一旦将一个实际问题表述为凸优化问题，大体上意味着相应问题已经得到彻底解决，这是非凸的优化问题所不具有的性质。从理论角度看，用凸优化模型对一般性非线性优化模型进行局部逼近，始终是研究非线性规划问题的主要途径，因此，通过学习凸优化理论，可以直接或间接地掌握数学规划领域几乎所有重要的理论结果。由于上述原因，对于涉足优化领域的人员，无论是理论研究还是实际应用，都应该对凸优化理论和方法有一定程度的了解。

本书内容非常丰富。理论部分由4章构成，不仅涵盖了凸优化的所有基本概念和主要结果，还详细介绍了几类基本的凸优化问题以及将特殊的优化问题表述为凸优化问题的变换方法，这些内容对灵活运用凸优化知识解决实际问题非常有用。应用部分由3章构成，分别介绍凸优化在解决逼近与拟合、统计估计和几何关系分析这三类实际问题中的应用。算法部分也由3章构成，依次介绍求解无约束凸优化模型、等式约束凸优化模型以及包含不等式约束的凸优化模型的经典数值方法，以及如何利用凸优化理论分析这些方法的收敛性质。通过阅读本书，能够对凸优化理论和方法建立完整的认识。

本书对每章内容都配备了大量习题，因此也非常适合用作教科书。实际上，该书多年来已在美国多所大学用于课堂教学，近两年也在清华大学自动化系用作相关研究生课程的主要教材。

本书前言及第1, 3, 5, 7章由许鋈翻译；第2, 4, 6, 8章由黄晓霖翻译；第9~11章及附录部分由王书宁翻译。全书由王书宁定稿。

感谢Stephen Boyd教授提供本书英文版本的电子文件，为翻译本书提供了极大的便利。

译者

2012年10月

前言

本书研究优化问题的一个重要分支：**凸优化**。事实上，最小二乘以及线性规划问题都属于凸优化问题。众所周知，关于最小二乘和线性规划问题的理论相当成熟，这两个问题出现在很多应用领域，均能很快地进行数值求解。本书的基本观点是，除了这两个问题以外，还有很多凸优化问题亦是如此。

尽管凸优化的研究已经持续了一个世纪左右，然而，最近一些相关的研究成果使得这一问题重新引起人们的关注。这当中首推对内点法的重新认识。内点法于 20 世纪 80 年代提出，本是用以求解线性规划问题，但是最近人们认识到，它亦可以被应用于求解凸优化问题。这些新的方法使得我们可以如求解线性规划一样有效求解一些特殊的凸优化问题，如半定规划以及二阶锥规划问题。

第二个相关的研究成果是人们发现凸优化问题（不仅仅是最小二乘和线性规划）在实践中的应用远远超乎人们的想象。从 20 世纪 90 年代开始，凸优化即被用在自动控制系统、估计和信号处理、通信网络、电路设计、数据分析及建模、统计和金融方面。此外，在组合优化以及全局优化方面，凸优化经常被用来估计最优值的界以及给出近似解。我们相信，还有很多其他凸优化的应用领域正在等待着人们去发现。

发现某个问题是凸优化问题或能将其描述为凸优化问题将会大有裨益。最本质的好处就是对此问题可以用内点法或者其他凸优化方法进行可靠且迅速的求解。这些求解方法可靠，足以嵌入于电脑辅助设计或分析工具，甚至用于实时响应系统或者自动控制系统。此外，将某个问题描述为凸优化问题还具有理论或概念上的优越性。例如，相应的对偶问题经常可以基于原问题给出有意义的解释，有时可导向有效的或分布式的求解原问题的方法。

我们认为，凸优化非常重要，任何从事计算数学的人至少需要对其有一定的了解。在我们看来，凸优化理所当然地是继近代线性代数（如最小二乘，奇异值）和线性规划之后的又一重要领域。

本书目的

对于很多一般性的优化方法，通常人们用它直截了当地试解待解的问题。凸优化就与此不同，只有我们知道待解的问题是凸的，它的优越性才可能完整地体现出来。当然，很多优化问题是非凸的，判断某个问题是否凸或者将某个问题表述为凸优化的形式是比较困难的。

本书的主要目的是帮助读者掌握应用凸优化方法的相关知识，即判断、描述

以及求解凸优化问题的技能和背景知识。

获取凸优化的相关应用知识对数学要求较高，对于主要关注应用的读者更是如此。根据我们的经验，对于电气工程以及计算科学的研究生来说，在这方面的投入会获得良好的、有时是丰厚的回报。

在此之前，有不少关于线性规划以及一般的非线性规划的书籍，这些书籍的侧重点在于问题的描述，建模以及应用。另有一些书籍主要讨论凸优化的理论，内点法以及复杂度分析。本书介于二者之间，介绍一般的凸优化理论，侧重于问题描述以及建模。

我们也要指出本书并不追求什么。它不是一本侧重于凸分析或者凸优化的数学知识的教材，已经有一些别的书籍涵盖了这些内容。此外，本书也不是凸优化算法的一个综述。我们只是挑选了一些较好的算法，介绍其简化了的或者是典型的形式（但是它们在实践中确实能发挥作用）。本书也不试图涵盖求解凸问题的内点法（或其他方法）的最新发展动态。虽然本书所提供的一些数值仿真实例经过了高度简化，但是我们认为，它们能够适应一些潜在用户的应用要求。并且，对于一些凸优化算法，本书详细地探讨了如何利用问题的结构使得求解更为迅速。对于所描述算法的复杂性理论，我们也只是以一种简单方式进行了介绍。然而，对于内点法的自和谐和复杂度分析的重要思想我们都有一定的介绍。

读者范围

对于在工作中需要用到数学优化，或者更一般地说，用到计算数学的科研人员、科学家以及工程师，本书较为适合。这些人群包括直接从事优化或者运筹学的科技工作者，亦包括一些工作在其他科学和工程领域但是需要借助数学优化工具的科技工作者，这些领域包括计算科学、经济学、金融、统计学、数据挖掘等。本书主要针对后者，即可能使用凸优化的科技工作者，而不是针对人数相对少很多的凸优化领域的专家。

在阅读本书之前，读者只需要掌握现代微积分和线性代数的相关知识。如果读者对一些基本的数学分析知识（如范数、收敛性、初等拓扑学）和基本的概率论有一定的了解，应能较好地理解本书的所有论证和讨论。当然，我们希望即使没有学过数学分析和概率论的读者也能够理解本书所有的基本思想和要点。此外，本书的正文以及附录部分包含了数值计算和优化方法需要的所有辅助资料，因此，读者并不需要事先具备这些知识。

使用本书作为教材

我们希望本书能够在不同的课程中作为基本教材或者是参考教材发挥它的作用。从 1995 年开始，我们即在 Stanford 和 UCLA 的一些研究生课程中使用本书的初稿，这些课程包括线性优化、非线性优化和凸优化（偏工程应用）。我们的经验表明，用一个研究生课程的四分之一时间即可以粗略讲授完本书的大部分内容，如果用一个月学

期的课程时间，讲课进度就可以比较从容，也可以增加更多的例子，并且可以更加详尽地讨论有关理论。若能用两个四分之一的研究生课程时间，就可以对线性规划和二次规划（对于以应用为目的的学生极为重要）这些基本内容进行较广泛的讨论，或者对学生布置更多的大练习。

本书可以作为线性优化、非线性优化等基础课的参考读物。对于涉及凸优化的应用领域如控制系统等课程，本书亦可以作为替换教材。此外，对于凸优化方面更关注理论的课程，本书可以作为辅助教材，它提供了一些简单的实际例子。

致谢

本书的完成历时将近十年。这十年中，我们收到了不少关于本书的反馈以及建议，这些建议来自我们的研究生、我们课程上的学生以及我们在 Stanford 和 UCLA 的同事等。篇幅有限，我们无法一一表达我们的感谢，仅列出下述名单，表达我们诚挚的谢意。A. Aggarwal, V. Balakrishnan, A. Bernard, B. Bray, R. Cottle, A. d'Aspremont, J. Dahl, J. Dattorro, D. Donoho, J. Doyle, L. El Ghaoui, P. Glynn, M. Grant, A. Hansson, T. Hastie, A. Lewis, M. Lobo, Z.-Q. Luo, M. Mesbahi, W. Naylor, P. Parrilo, I. Pressman, R. Tibshirani, B. Van Roy, L. Xiao 和 Y. Ye. 我们要感谢 J. Jalden 以及 A. d'Aspremont 在时间序列分析 §6.5.4 中所提供的例子，§6.5.5 中的界定顾客喜好的例子也由他们提供。此外，感谢 P. Parrilo 对习题 4.4 和习题 4.56 所提供的建议。

我们还要特别感谢两个人。Arkadi Nemirovski 引发了我们对凸优化的兴趣并且鼓励我们撰写本书。而 Kishan Baheti 对本书的完成也发挥了极大的作用。早在 1994 年的时候，他就鼓励我们以凸优化在实际工程中的应用为题申请美国科学基金会的科研课程基金，本书可以认为是当年的基金成果，虽然在时间上可能有所滞后。

Stephen Boyd
Lieven Vandenberghe

Stanford, California
Los Angeles, California

2003 年 7 月

目录

1	引言	1
1.1	数学优化	1
1.2	最小二乘和线性规划	3
1.3	凸优化	6
1.4	非线性优化	8
1.5	本书主要内容	10
1.6	符号	12
	参考文献	13
1	理论	17
2	凸集	19
2.1	仿射集合和凸集	19
2.2	重要的例子	24
2.3	保凸运算	31
2.4	广义不等式	38
2.5	分离与支撑超平面	42
2.6	对偶锥与广义不等式	46
	参考文献	52
	习题	53
3	凸函数	61
3.1	基本性质和例子	61
3.2	保凸运算	73
3.3	共轭函数	85
3.4	拟凸函数	90

3.5	对数-凹函数和对数-凸函数	98
3.6	关于广义不等式的凸性	102
	参考文献	106
	习题	106
4	凸优化问题	121
4.1	优化问题	121
4.2	凸优化	130
4.3	线性规划问题	139
4.4	二次优化问题	145
4.5	几何规划	153
4.6	广义不等式约束	160
4.7	向量优化	167
	参考文献	179
	习题	180
5	对偶	207
5.1	Lagrange 对偶函数	207
5.2	Lagrange 对偶问题	215
5.3	几何解释	224
5.4	鞍点解释	229
5.5	最优性条件	233
5.6	扰动及灵敏度分析	241
5.7	例子	245
5.8	择一定理	250
5.9	广义不等式	256
	参考文献	264
	习题	265
II	应用	283
6	逼近与拟合	285
6.1	范数逼近	285
6.2	最小范数问题	295
6.3	正则化逼近	297

6.4	鲁棒逼近	307
6.5	函数拟合与插值	314
	参考文献	329
	习题	329
7	统计估计	337
7.1	参数分布估计	337
7.2	非参数分布估计	345
7.3	最优检测器设计及假设检验	350
7.4	Chebyshev 界和 Chernoff 界	360
7.5	实验设计	370
	参考文献	376
	习题	377
8	几何问题	381
8.1	向集合投影	381
8.2	集合间的距离	386
8.3	Euclid 距离和角度问题	389
8.4	极值体积椭球	394
8.5	中心	400
8.6	分类	406
8.7	布局与定位	414
8.8	平面布置	420
	参考文献	426
	习题	427
III	算法	435
9	无约束优化	437
9.1	无约束优化问题	437
9.2	下降方法	443
9.3	梯度下降方法	445
9.4	最速下降方法	454
9.5	Newton 方法	461
9.6	自和谐	473

9.7 实现	484
参考文献	489
习题	489
10 等式约束优化	497
10.1 等式约束优化问题	497
10.2 等式约束的 Newton 方法	501
10.3 不可行初始点的 Newton 方法	507
10.4 实现	520
参考文献	531
习题	531
11 内点法	535
11.1 不等式约束的极小化问题	535
11.2 对数障碍函数和中心路径	536
11.3 障碍方法	542
11.4 可行性和阶段 1 方法	552
11.5 自和谐条件下的复杂性分析	558
11.6 广义不等式问题	568
11.7 原对偶内点法	580
11.8 实现	587
参考文献	592
习题	593
附录	603
A 有关的数学知识	605
A.1 范数	605
A.2 分析	609
A.3 函数	610
A.4 导数	612
A.5 线性代数	617
参考文献	623

B	双二次函数的问题	624
B.1	单约束二次优化	624
B.2	S-程序	626
B.3	双对称矩阵的数值场	627
B.4	强对偶结果的证明	629
	参考文献	630
C	有关的数值线性代数知识	631
C.1	矩阵结构与算法复杂性	631
C.2	求解已经因式分解的矩阵的线性方程组	634
C.3	LU, Cholesky 和 LDL^T 因式分解	637
C.4	分块消元和 Schur 补	642
C.5	求解不确定线性方程组	650
	参考文献	653
	参考文献	654
	符号	670
	索引	673

1 引言

引言将对数学优化做一个简单的回顾，我们将主要关注凸优化在数学优化中的特殊地位。对于在引言中出现的一些概念，后面将会更加正式以及详细地进行定义。

1.1 数学优化

数学优化问题或者说优化问题可以写成如下形式

$$\begin{aligned} & \text{minimize} && f_0(x) \\ & \text{subject to} && f_i(x) \leq b_i, \quad i = 1, \dots, m. \end{aligned} \tag{1.1}$$

这里，向量 $x = (x_1, \dots, x_n)$ 称为问题的优化变量，函数 $f_0: \mathbf{R}^n \rightarrow \mathbf{R}$ 称为目标函数，函数 $f_i: \mathbf{R}^n \rightarrow \mathbf{R}$, $i = 1, \dots, m$, 被称为（不等式）约束函数，常数 b_1, \dots, b_m 称为约束上限或者约束边界。如果在所有满足约束的向量中向量 x^* 对应的目标函数值最小：即对于任意满足约束 $f_1(z) \leq b_1, \dots, f_m(z) \leq b_m$ 的向量 z ，有 $f_0(z) \geq f_0(x^*)$ ，那么称 x^* 为问题 (1.1) 的最优解或者解。

我们一般考虑具有特殊形式的目标函数和约束函数的优化问题。线性规划即是其中重要的一类。若优化问题 (1.1) 中的目标函数和约束函数 f_0, \dots, f_m 都是线性函数，即对任意的 $x, y \in \mathbf{R}^n$ 和 $\alpha, \beta \in \mathbf{R}$ 有

$$f_i(\alpha x + \beta y) = \alpha f_i(x) + \beta f_i(y), \tag{1.2}$$

则此优化问题称为线性规划。若优化问题不是线性的，则称之为非线性规划。

本书讨论一类优化问题，即凸优化问题。凸优化问题中的目标函数和约束函数都是凸函数，即对于任意 $x, y \in \mathbf{R}^n$ ，任意 $\alpha, \beta \in \mathbf{R}$ 且满足 $\alpha + \beta = 1$, $\alpha \geq 0$, $\beta \geq 0$ ，下述不等式成立

$$f_i(\alpha x + \beta y) \leq \alpha f_i(x) + \beta f_i(y). \tag{1.3}$$

比较式 (1.3) 和式 (1.2)，可以看出凸性是比较线性更为一般的性质：线性函数需要严格满足等式，而凸函数仅仅需要在 α 和 β 取特定值的情况下满足不等式。因此线性规划问题也是凸优化问题，可以将凸优化看成是线性规划的扩展。

1.1.1 应用

优化问题 (1.1) 可以看成在向量空间 \mathbf{R}^n 的一集备选解中选择最好的解。用 x 表示备选解, $f_i(x) \leq b_i$ 表示 x 必须满足的条件, 目标函数 $f_0(x)$ 表示选择 x 的成本 (同理也可以认为 $-f_0(x)$ 表示选择 x 的效益或者效用)。优化问题 (1.1) 的解即为满足约束条件的所有备选解中成本最小 (或者效用最大) 的那个解。

以**投资组合优化**为例, 我们寻求一个最佳投资方案将资本在 n 种资产中进行分配。变量 x_i 表示投资在第 i 种资产上的资本, 那么向量 $x \in \mathbf{R}^n$ 则描述了资本在各个资产上的分配情况。约束条件可能包括对投资预算的限制 (即总投资额的限制), 每一份投资额非负 (此处假设不允许空头) 以及期望收益必须大于最小可接受收益值。目标函数或者说成本函数可能是对总的风险值或者投资回报方差的一个度量。在此意义下, 优化问题 (1.1) 即为在所有可能的投资组合中选择满足所有约束条件且风险最小的那个组合。

另一个例子是电子设计中的**器件尺寸问题**, 即在电子电路中设计每个器件的长度和宽度。此时优化变量表示器件的长度和宽度。约束条件表征了多种工程上需要满足的要求, 如生产过程对器件尺寸的要求, 电路在特定速度稳定运行对时间的要求, 或者是电路的总面积的限制。器件尺寸设计问题中一个比较常见的目标函数是电路的总功耗。因此, 优化问题 (1.1) 此时变为设计合适的电路器件尺寸, 使之满足设计要求 (制造要求、时间要求以及面积要求) 并且最为节能。

在**数据拟合**中, 人们需要在一族候选模型中选择最符合观测数据与先验知识的模型。此时, 变量为模型中的参数, 约束可以是先验知识以及参数限制 (比如说非负性)。目标函数可能是与真实模型的偏差或者是观测数据与估计模型的预测值之间的偏差, 也有可能是参数值的似然度和置信度的统计估计。优化问题 (1.1) 此时即为寻找合适的模型参数值, 使之符合先验知识, 且与真实模型之间的偏差或者预测值与观测值之间的偏差最小 (或者在统计意义上更加相似)。

大量涉及决策 (或系统设计、分析与操作) 的实际问题可以表示成数学优化问题, 或者数学优化问题的变化形式如多目标优化问题。事实上, 数学优化已经成为很多领域的重要工具, 它被广泛地应用于工程、电路自动设计、自动控制系统以及土木、化工、机械、航空工程中出现的最优设计问题。此外, 网络设计和操作, 金融, 供应链管理, 调度等很多领域的问题都需要用到优化。反映这类应用的列表还在稳定地扩展。

在上述大部分应用中, 数学优化都是作为一个辅助工具, 协助决策者, 系统设计人员或系统操作员根据需要监控系统的运行过程, 检查结果以及修改问题 (或者方法)。这些人员根据优化问题的结果采取相应的措施, 如购买或出售资产以得到最佳投资。

最近出现的一些现象使得数学优化在其他一些领域中的应用成为可能。随着越来越多的计算机嵌入式产品的问世, **嵌入式优化**的发展势头迅猛。在这些嵌入式应用中,

优化的目的是在没有（或者极少）人为干预的条件下实时做出选择甚至实时执行动作。在一些应用领域，传统自动控制系统以及嵌入式优化已经被有机地融合在一起；而在其他一些领域关于这种结合的尝试还处在一个初始阶段。嵌入式实时优化同样带来一些新的挑战：优化结果必须非常可靠且必须在给定的时间（和存储量）内解决问题。

1.1.2 求解优化问题

某类优化问题的求解方法是指（以给定精度）求解此类优化问题中的某一实例的算法。从 20 世纪 40 年代后期开始，学者们致力于设计算法求解多类优化问题，分析这些算法的性质以及进行软件实现。不同算法的有效性，即我们用之求解优化问题 (1.1) 的能力，是大不相同的。它取决于多方面的因素，如目标函数和约束函数的形式，优化问题所包含的变量和约束的个数以及一些特殊的结构如稀疏结构（如果某个问题中每个约束函数仅仅取决于为数不多的几个变量，那么此问题称为稀疏的）。

即使目标函数和约束函数是光滑的（如多项式），一般形式的优化问题 (1.1) 仍然很难求解。因此，求解一般形式的问题是需要付出一些代价的，如需要较长的计算时间或者可能找不到解。§1.4 讨论了一些这样的算法。

然而，并不是所有的优化问题都难以求解。对于一些特殊的优化问题，存在一些有效的算法，这些算法对含有成百上千变量和约束的大型问题甚至都有效。两类重要且广为人知的例子是最小二乘问题和线性规划问题，随后的 §1.2 将描述这两个问题（第 4 章中将予以详细讨论）。鲜为人知的是，还有一类问题，存在有效的求解算法可以如最小二乘以及线性规划问题一样进行有效求解。这类问题即为凸优化问题，一些大型的凸优化问题甚至都能被可靠求解。

1.2 最小二乘和线性规划

本节介绍广为人知且应用广泛的两类特殊的凸优化问题，最小二乘和线性规划（第 4 章将详细讨论这两类问题）。

1.2.1 最小二乘问题

最小二乘问题是这样一类优化问题，它没有约束条件（即 $m = 0$ ），目标函数是若干项的平方和，每一项具有形式 $a_i^T x - b_i$ ，具体形式如下

$$\text{minimize } f_0(x) = \|Ax - b\|_2^2 = \sum_{i=1}^k (a_i^T x - b_i)^2. \quad (1.4)$$

其中， $A \in \mathbf{R}^{k \times n}$ ($k \geq n$)， a_i^T 是矩阵 A 的行向量，向量 $x \in \mathbf{R}^n$ 是优化变量。

求解最小二乘问题

最小二乘问题 (1.4) 的求解可以简化为求解一组线性方程

$$(A^T A)x = A^T b,$$

因此可得解析解 $x = (A^T A)^{-1} A^T b$ 。对于最小二乘问题，有不少好的求解算法（和软件实现），这些算法的精度和可靠性都很高。最小二乘问题可以在有限时间内进行求解，此时间和 $n^2 k$ 近似成正比，且比例系数已知。现有的台式计算机可以在几秒之内解决包含数百个变量、数千个求和项的最小二乘问题；当然，更为先进的计算机可以解决更大规模的问题或者以更快的速度解决相同规模的问题（此外，根据摩尔定律，以后的求解时间还会指数下降）。求解最小二乘问题的算法和软件可靠，足以满足嵌入式优化的要求。

很多时候，若最小二乘问题中的系数矩阵 A 具有某些特殊结构，我们甚至可以解决更大规模的此类问题。例如，假设矩阵 A 是稀疏的，即矩阵 A 中含有的非零元素的个数远远少于 kn ，我们可以更快求解此最小二乘问题，所需时间的数量级远远小于 $n^2 k$ 。现有的台式计算机在一分钟左右的时间内可以处理大规模的稀疏最小二乘问题，这些问题中所含的变量个数甚至多达数万，所含的求和项也可达到数十万（具体时间取决于不同的稀疏类型）。

当最小二乘问题的规模极大（比如说包含上百万的变量）或者需要在给定时间内进行实时求解，此时求解最小二乘问题就有一定的难度。然而，在大部分的应用中，已有的方法还是相当有效和可靠的。事实上，我们可以认为，最小二乘问题的求解是一项（成熟的）技术（除非问题超出目前可解的范围），对于很多人，即使不知道，也无需知道这门技术的细节，仍然可以可靠地应用。

使用最小二乘

最小二乘问题是回归分析，最优控制以及很多参数估计和数据拟合方法的基础。最小二乘问题有很多统计意义，例如，给定包含高斯噪声的线性测量值时，向量 x 的最大似然估计即等价于最小二乘问题的解。

判别一个优化问题是否是最小二乘问题非常简单；只需要检验目标函数是否是二次函数（然后检验此二次函数是否半正定）。基本最小二乘问题只有一个简单固定的表达式，因此，人们提出了一些标准的技术，使得最小二乘问题在实际应用中更为灵活。

在加权最小二乘问题中，我们最小化加权的最小二乘成本

$$\sum_{i=1}^k w_i (a_i^T x - b_i)^2,$$

其中，加权系数 w_1, \dots, w_k 均大于零（加权最小二乘问题可以很方便地转化为标准的最小二乘问题进行求解）。在加权最小二乘问题中，加权系数 w_i 反映了求和项 $a_i^T x - b_i$ 的重要程度或者说是解的影响程度。在统计应用中，当给定的线性观测值包含不同方差的噪声时，我们用加权最小二乘来估计向量 x 。

正则化是解决最小二乘问题的另一个技术，它通过在成本函数中增加一些多余的项来实现。一个最简单的形式是在成本函数中增加一项和变量平方成正比的项

$$\sum_{i=1}^k (a_i^T x - b_i)^2 + \rho \sum_{i=1}^n x_i^2,$$

这里， $\rho > 0$ 。（此问题亦可以转化为标准最小二乘问题。）当 x 的值较大时，增加的项对其施加一个惩罚，其得到的解比仅优化第一项时更加切合实际。参数 ρ 的选择取决于使用者，选择原则是使原始目标函数值尽可能小而同时保证 $\sum_{i=1}^n x_i^2$ 的值不能太大，在二者之间取得一个较好的平衡。在统计估计中，当待估计向量 x 的分布预先知道时，可以采用正则化方法。

加权最小二乘和正则化在第 6 章会进行详细介绍；第 7 章给出了它们在统计学上的意义。

1.2.2 线性规划

另一类较为重要的优化问题是线性规划，其目标函数和所有的约束函数均为线性函数，线性规划问题可以表述如下

$$\begin{aligned} & \text{minimize} && c^T x \\ & \text{subject to} && a_i^T x \leq b_i, \quad i = 1, \dots, m. \end{aligned} \quad (1.5)$$

其中，向量 $c, a_1, \dots, a_m \in \mathbf{R}^n$ ， $b_1, \dots, b_m \in \mathbf{R}$ 是问题参数，它们决定目标函数和约束函数。

求解线性规划

虽然线性规划问题的解并没有一个简单的解析表达形式（和最小二乘问题不同），然而，存在很多非常有效的求解线性规划问题的方法，这当中包括 Dantzig 的单纯形法以及最近发展起来的内点法。本书后面将专门介绍内点法。此外，我们也不能给出解决一个线性规划问题所需要的算术运算的确切次数，但对于给定的求解精度，可以给出采用内点法时所需要的算术运算次数的严格上界。其求解复杂度实际正比于 $n^2 m$ （假设 $m \geq n$ ），但是比例系数不好确定，而最小二乘方法中比例系数较易确定。尽管与最小二乘问题的算法相比可靠性略逊一筹，但求解线性规划问题的算法还是相当可靠的，可以在大约几秒钟的时间内利用小型台式计算机处理含有数百变量、数千约束条件的线性规划问题；如果线性规划问题是稀疏的，或者具有其他有利于运算的结构，甚至可以解决包含数万或者数十万变量的问题。

和最小二乘问题一样，处理极大规模的线性规划问题或者在很短时间内实时解决线性规划问题还是具有一定难度的。但是，和最小二乘问题的情况类似，我们可以说求解（大部分）线性规划问题是一项成熟的技术。线性规划的求解程序可以（并已经）嵌入到很多工具箱和应用软件中。