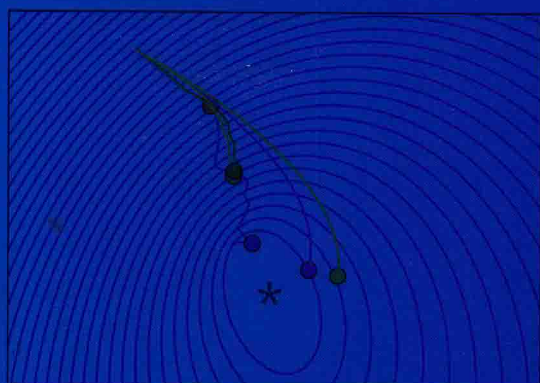
工科研究生数学类基础课程应用系列丛书

现代优化理论与方法

（下册）

黄庆道　吕显瑞　李晓峰　王彩玲　编



科学出版社

工科研究生数学类基础课程应用系列丛书

现代优化理论与方法

(下册)

黄庆道　吕显瑞　李晓峰　王彩玲　编

科学出版社
北京

内 容 简 介

本书分上、下两册，共 11 章，包括最优化问题、线性规划、非线性规划问题、多目标规划、全局最优化问题、二次规划、整数规划、动态规划及优化求解的软件实现等问题.

本书可作为最优化及相关专业的研究生教材和高年级本科生的选修教材，也可供从事相关专业的科研人员和工程技术人员参考.

图书在版编目(CIP)数据

现代优化理论与方法. 下册/黄庆道等编. —北京：科学出版社，2017.6
(工科研究生数学类基础课程应用系列丛书)
ISBN 978-7-03-049673-7

Ⅰ. ①现… Ⅱ. ①黄… Ⅲ. ① 最佳化–数学理论–研究生–教材 Ⅳ. ①O224

中国版本图书馆 CIP 数据核字(2016) 第 201916 号

责任编辑：张中兴 / 责任校对：彭 涛
责任印制：白 洋 / 封面设计：迷底书装

科学出版社 出版
北京东黄城根北街 16 号
邮政编码：100717
http://www.sciencep.com

三河市书文印刷有限公司 印刷
科学出版社发行 各地新华书店经销
*
2017 年 6 月第 一 版 开本：720 × 1000 1/16
2017 年 6 月第一次印刷 印张：8 1/2
字数：172 000

定价：29.00 元
(如有印装质量问题，我社负责调换)

《工科研究生数学类基础课程应用系列丛书》
编委会名单

序　言

研究生公共数学系列教材是根据国家教育部关于研究生培养指导规划和目标、结合当前研究生教育改革的实际情况、借鉴国内外研究生教育的最新研究成果、旨在规范和加强研究生公共基础课教学的一套研究生公共数学系列教材. 本系列教材经过对研究生公共数学课程整合、优化, 共编写教材 13 册, 具体包括:《现代分析基础》(上、下)、《代数学基础》(上、下)、《现代统计学基础》(上、下)、《现代微分方程概论》(上、下)、《现代计算方法》(上、下)、《现代优化理论与方法》(上、下)、《应用泛函分析》. 其中上册为非数学类硕士研究生教材, 下册为非数学类博士研究生教材.

本系列教材的编写体现了时代的特征, 本着加强基础、淡化证明、强调应用的原则, 力争做到科学性、系统性和实用性的统一, 着眼于传授教学知识和培养学生数学素养的高度结合.

本系列教材吸取了国内外同类教材的精华, 参考了近年来出版的一些新教材, 结合了当前研究生公共数学教学改革的实际, 特别是综合性大学非数学类研究生公共数学的实际需求.

本系列教材体例科学、结构合理、内容经典而追求创新, 既是作者多年教学经验的总结, 又是作者长期教学研究和科学研究成果的体现. 每章后面配备巩固基本概念、基本理论、基本运算的基本题目, 又有提高学生抽象思维、逻辑推理和综合运用基础知识解题的提高题目, 为学生掌握教材基本内容, 运用教材基本知识开发创新思维提供了可行条件.

本系列教材适用面广、涉及专业全、教学内容新, 可作为综合性大学非数学专业研究生公共数学教材和数学参考书, 在教材体系与内容的编排上认真考虑了不同专业, 不同学时的授课对象的需求, 可选择不同的教学模块, 以满足广大读者的实际需要.

本系列教材的编写过程中, 得到了吉林大学研究生院、吉林大学数学学院和数学研究所的大力支持, 也得到了科学出版社的领导和编辑们的鼎力帮助, 在此一并致谢.

由于编者水平有限, 书中的不妥之处在所难免, 恳请广大读者批评指正, 以期不断完善.

研究生公共数学课程系列教材编委会

2015 年 3 月于长春

前　言

本课程属于非数学类研究生数学公共基础课程之一，最优化方法是从所有可能方案中选择最合理的方案以达到最优目标的科学. 随着电子计算机的普遍应用而迅猛发展，已广泛应用于国民经济各部门和科学技术的各个领域中. 因此，学习和掌握最优化的基本理论和方法，对于将来从事工程技术工作的工科研究生来说是必不可少的. 本门课程旨在讲授最优化的基本理论和方法，要求通过本课程的学习，具有应用最优化方法解决一些实际问题的初步技能，并为以后的学习和工作做必要的准备.

根据工科研究生课程指导委员会制定的“工科研究生最优化方法课程教学基本要求”，结合吉林大学为本校硕士研究生和本科生编写的最优化理论与方法教材及多年来教学实践的体会，我们选取了线性规划、非线性规划、多目标规划、全局最优化和现代优化理论五部分，即全书 1—7 章，作为优化基本理论上册内容.

为了进一步优化理论的学习，我们又选取了二次规划、整数规划、动态规划和优化求解的软件实现的内容，即全书 8—11 章作为本书的下册内容. 每一部分内容着重阐明基本理论与基本方法，以便给读者在该领域的深入学习和研究打下良好基础，对于一些证明较冗长和复杂的定理，我们只给出定理的内容，证明从略.

本书力求深入浅出，通俗易懂，学过高等数学和线性代数的读者均能学习. 本书既可作为工科研究生和高年级本科生学习本门课程的教材，也可以作为从事应用数学、管理学、系统工程及工程设计方面的广大科技工作者的参考书.

由于编者水平有限，缺点和疏漏在所难免，敬请读者予以批评指正.

黄庆道
2016 年 3 月

目　录

的非零向量 d 都有

$$d^{\mathrm{T}}Hd \geqslant 0, \tag{8.13}$$

则 x^* 必是问题 (8.1)—(8.3) 的局部严格极小点.

定理 8.1.3　设 x^* 是二次规划问题 (8.1)—(8.3) 的可行点, 则 x^* 是一局部极小点当且仅当存在乘子 $\lambda^* = (\lambda_1^*, \cdots, \lambda_m^*)$ 使得 (8.4)—(8.6) 成立而且对一切满足 (8.10)—(8.12) 的向量 d 都有

$$d^{\mathrm{T}}Hd \geqslant 0. \tag{8.14}$$

证明　设 x^* 是一局部极小点, 由定理 8.1.1 知存在乘子 λ^* 使得 (8.1)—(8.3) 成立. 设 d 是任何一个满足 (8.10)—(8.12) 的非零向量. 显然, 对充分小的 $t > 0$ 有

$$x^* + td \in X. \tag{8.15}$$

于是, 由 d 的定义,

$$\begin{aligned} Q(x^*) \leqslant Q(x^* + td) &= Q(x^*) + td^{\mathrm{T}}[Hx^* + g] + \frac{1}{2}t^2 d^{\mathrm{T}}Hd \\ &= Q(x^*) + t\sum_{i=1}^{m} \lambda_i^* a_i^{\mathrm{T}} d + \frac{1}{2}t^2 d^{\mathrm{T}}Hd \\ &= Q(x^*) + \frac{1}{2}t^2 d^{\mathrm{T}}Hd \end{aligned} \tag{8.16}$$

对所有充分小的 $t > 0$ 成立, 故知 (8.14) 式成立. 由于 d 的任意性, 所以对一切满足于 (8.10)—(8.12) 的向量 d 都有 (8.14).

反之, 设存在 $\lambda^* = (\lambda_1^*, \cdots, \lambda_m^*)$ 使得 (8.4)—(8.6) 成立, 而且对一切满足于 (8.10)—(8.12) 的向量 d 都有 (8.14) 式成立. 如果 x^* 不是一局部极小点, 则必存在 $\delta_k > 0$, d_k 使得

$$x^* + \delta_k d_k \in X, \tag{8.17}$$

$$Q(x^* + \delta_k d_k) < Q(x^*), \tag{8.18}$$

而且 $\delta_k \to 0, d - k \to \bar{d}$. 考虑 Lagrange 函数

$$L(x, \lambda^*) = Q(x) - \sum_{i=1}^{m} \lambda_i^* (a_i^{\mathrm{T}} x - b_i). \tag{8.19}$$

由于 $L(x, \lambda^*)$ 是关于 x 的二次函数且有

$$\nabla_x L(x^*, \lambda^*) = 0, \tag{8.20}$$

$$\nabla_{xx} L(x^*, \lambda^*) = H, \tag{8.21}$$

$$
\begin{aligned}
L(x^*+\delta_k d_k,\lambda^*) &= Q(x^*+\delta_k d_k)-\sum_{i\in I}\lambda_i^*\delta_k a_i^{\mathrm{T}}d_k \\
&\leqslant Q(x^*)-\sum_{i\in I}\lambda_i^*\delta_k a_i^{\mathrm{T}}d_k. \qquad (8.22)
\end{aligned}
$$

由 (8.20) 和 (8.22) 可知 $\bar{d}$ 满足 (8.10)—(8.12). 令矩阵 $\bar{A}$ 是由 $a_i(i\in E,\lambda_i^*>0,i\in I)$ 组成的. 定义

$$\bar{d_k}=-(\bar{A}^{\mathrm{T}})+\bar{A}^{\mathrm{T}}d_k, \qquad (8.23)$$

$$\widehat{d_k}=d_k+\bar{d_k}. \qquad (8.24)$$

由 (8.22) 可知

$$\|\bar{d_k}\|\to 0. \qquad (8.25)$$

由 (8.20) 和 (8.21) 可知

$$
\begin{aligned}
L(x^*+\delta_k d_k,\lambda^*) &= L(x^*,\lambda^*)+\frac{1}{2}\delta_k^2[(\widehat{d_k}-\bar{d_k})^{\mathrm{T}}H(\widehat{d_k}-\bar{d_k})] \\
&\geqslant L(x^*,\lambda^*)+O(\|\bar{d_k}\|_2\delta_k^2) \\
&= Q(x^*)+O(\delta_k^2\|\bar{A}^{\mathrm{T}}d_k\|_2). \qquad (8.26)
\end{aligned}
$$

从 (8.22) 和 (8.26) 可得到

$$(\min_{\substack{i\in I\\ \lambda_i^*>0}}\lambda_i^*)\delta_k\|\bar{A}^{\mathrm{T}}d_k\|_2\leqslant O(\delta_k^2\|\bar{A}^{\mathrm{T}}d_k\|_2). \qquad (8.27)$$

这与 $\delta_k\to 0$ 相矛盾. 矛盾说明 x^* 必是一局部极小点. □

由于问题的特殊形式, 求解二次规划的 K-T 点等价于寻求 $x^*\in\mathbf{R}^n,\lambda^*\in\mathbf{R}^m$ 使得线性系统 (8.2)—(8.3), (8.4), (8.6) 满足而且线性互补条件 (8.5) 也成立.

如果 H 是 (正定) 半正定矩阵, (8.1) 中的目标函数是 (严格) 凸函数, 这时问题 (8.1)—(8.3) 被称为 (严格) 凸的二次规划问题. 对于二次规划, 可行域只要不空就必定是凸集, 所以但当目标函数是凸函数时, 任何 K-T 点必为二次规划的全局极小点.

定理 8.1.4 设 H 是半正定矩阵, 则 x^* 是二次规划问题 (8.2)—(8.3) 的全局极小点当且仅当它是一个局部极小点, 也当且仅当它是一个 K-T 点.

所以, 当 H 是半正定时, 求解 (8.2)—(8.3) 等价于求解 $(x,\lambda)\in\mathbf{R}^{m+n}$ 使得

$$g+Hx=A\lambda, \qquad (8.28)$$

$$a_i^{\mathrm{T}}x = b - i, \quad i \in E, \tag{8.29}$$

$$a_i^{\mathrm{T}}x \geqslant b - i, \quad i \in I, \tag{8.30}$$

$$\lambda_i[a_i^{\mathrm{T}}x - b_i] = 0, \quad i \in I, \tag{8.31}$$

$$\lambda_i \geqslant 0, \quad i \in I \tag{8.32}$$

成立, 其中 $I = m_e + 1, \cdots, m, \lambda = \lambda_1, \cdots, \lambda_m$ 以及

$$A = [a_1, \cdots, a_m]. \tag{8.33}$$

这一等价关系对推导凸二次规划的对偶规划问题是十分重要的.

8.2 对偶性质

假定 H 是正定矩阵, 由上节的结果可知二次规划 (8.1)—(8.3) 等价于 (8.28)—(8.32). 记

$$y = A\lambda - g, \tag{8.34}$$

$$t_i = a_i^{\mathrm{T}}x - b_i, \quad i \in I. \tag{8.35}$$

则 (8.28)—(8.32) 可写成下列形式

$$\begin{bmatrix} -b \\ H^{-1}y \end{bmatrix} = \begin{bmatrix} -A^{\mathrm{T}} \\ I \end{bmatrix} + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ t_{m_e+1} \\ \vdots \\ t_m \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \tag{8.36}$$

$$A\lambda - y = g, \tag{8.37}$$

$$\lambda_i \geqslant 0, \quad i \in I, \tag{8.38}$$

$$t_i\lambda_i = 0, \quad i \in I, \tag{8.39}$$

$$t_i \geqslant 0, \quad i \in I. \tag{8.40}$$

由定理 8.1.3 可知 (8.36)—(8.40) 等价于

$$\max \quad b^{\mathrm{T}}\lambda - \frac{1}{2}y^{\mathrm{T}}H^{-1}y = \bar{Q}(\lambda, y), \tag{8.41}$$

$$\text{s.t.} \quad A\lambda - y = g, \tag{8.42}$$

$$\lambda_i \geqslant 0, \quad i \in I. \tag{8.43}$$

由于问题 (8.41)—(8.43) 与问题 (8.1)—(8.3) 等价, 称 (8.41)—(8.43) 为 (8.1)—(8.3) 的对偶问题, 称 (8.1)—(8.3) 为原始问题. 利用 (8.34), 可将问题 (8.41)—(8.43) 简化成如下形式

$$\min_{\lambda \in \mathbf{R}^m} -(b + A^{\mathrm{T}}H^{-1}g)^{\mathrm{T}}\lambda + \frac{1}{2}\lambda^{\mathrm{T}}(A^{\mathrm{T}}H^{-1}A)\lambda, \tag{8.44}$$

$$\text{s.t.}\lambda_i \geqslant 0,\ i \in I. \tag{8.45}$$

假定 (λ, y) 是对偶问题 (8.41)—(8.43) 的可行点, x 是原始问题 (8.1)—(8.3) 的可行点, 则有

$$\begin{aligned} Q(x) - \bar{Q}(\lambda, y) &= x^{\mathrm{T}}[A\lambda - y] + \frac{1}{2}x^{\mathrm{T}}Hx \\ &\quad - \left[\lambda^{\mathrm{T}}A^{\mathrm{T}}x - \sum_{i \in I}\lambda_i t_i - \frac{1}{2}y^{\mathrm{T}}H^{-1}y\right] \\ &= \sum_{i \in I}\lambda_i t_i + \frac{1}{2}[x^{\mathrm{T}}Hx + y^{\mathrm{T}}H^{-1}y - 2x^{\mathrm{T}}y], \end{aligned} \tag{8.46}$$

其中 t_i 由 (8.35) 定义. 由于 H 正定, 显然有

$$Q(x) \geqslant \bar{Q}(\lambda, y). \tag{8.47}$$

从 (8.46) 式还看出, (8.47) 式两边相等当且仅当

$$\sum_{i \in I}\lambda_i(a_i^{\mathrm{T}}x - b_i) = 0, \tag{8.48}$$

$$x = H^{-1}y. \tag{8.49}$$

(8.49) 等价于 (8.28), 因为 x 是可行点, (8.48) 与 (8.31) 等价. 于是我们已经证明了下面的定理.

定理 8.2.1 设 H 正定, 如果原始问题有可行点, 则 $x^* \in X$ 是问题 (8.1)—(8.3) 的解当且仅当存在 (λ^*, y^*) 是对偶问题 (8.41)—(8.43) 之解且 $x^* = H^{-1}y^*$ 以及 λ^* 是原问题在 x^* 处的 Lagrange 乘子.

对于原始问题无可行点的情形, 我们有下列结果.

定理 8.2.2 设 H 正定, 则原始问题有可行点当且仅当对偶问题无界.

证明 如果原始问题有可行点, 由 (8.47) 式可知对偶问题的目标函数在满足 (8.42)—(8.43) 的集合上一致有上界.

现假设原始问题无可行点, 于是

$$(a_i^{\mathrm{T}}, b_i)\widetilde{x} = 0, \quad i \in E, \tag{8.50}$$

$$(a_i^{\mathrm{T}}, b_i)\widetilde{x} \geqslant 0, \quad i \in I, \tag{8.51}$$

$$(0, \cdots, 0, 1)\widetilde{x} < 0. \tag{8.52}$$

在 $\widetilde{x} \in \mathbf{R}^{n+1}$ 上无解. 由 Farkas 引理即知存在 $\bar{\lambda}_i (i = 1, \cdots, m)$ 使得

$$\sum_{i=1}^{m} \bar{\lambda}_i a_i = 0, \tag{8.53}$$

$$\sum_{i=1}^{m} \bar{\lambda}_i b_i = 1, \tag{8.54}$$

$$\bar{\lambda}_i \geqslant 0, \quad i \in I. \tag{8.55}$$

令 $\lambda_i = t\bar{\lambda}_i, y = -g$, 当 $t \to +\infty$ 时有

$$\bar{Q}(\lambda, y) = t \to +\infty.$$

而且对一切 $t > 0, \lambda = (t\bar{\lambda}_1, \cdots, t\bar{\lambda}_m)$ 和 $y = -g$ 满足约束条件 (8.42) 和 (8.43). 所以对偶问题无界. □

原始问题的 Lagrange 函数

$$L(x, \lambda) = Q(x) - \sum_{i=1}^{m} \lambda_i (a_i^{\mathrm{T}} x - b_i) \tag{8.56}$$

与对偶问题也是有着密切联系的. 不难看出, 求解 (8.28)—(8.32) 等价于求函数 $L(x, \lambda)$ 在区域 $\{(x, \lambda) | \lambda_i \geqslant 0, \ i \in I\}$ 上的稳定点. 由于 $L(x, \lambda)$ 的 Hesse 阵为

$$\nabla^2 L(x, \lambda) = \begin{bmatrix} H & -A \\ -A^{\mathrm{T}} & 0 \end{bmatrix}, \tag{8.57}$$

利用恒等式

$$\begin{bmatrix} I & 0 \\ A^{\mathrm{T}} H^{-1} & I \end{bmatrix} \nabla^2 L(x, \lambda) \begin{bmatrix} I & H^{-1}A \\ 0 & I \end{bmatrix} = \begin{bmatrix} H & 0 \\ 0 & A^{\mathrm{T}} H^{-1} A \end{bmatrix} \tag{8.58}$$

可知 $\nabla^2 L$ 恰恰有 n 个正特征值, 而且它的负特征值的个数正好为 A 的秩. 所以, $L(x,\lambda)$ 的稳定点一般是一个鞍点.

事实上, 对任何 $x\in X$ 有

$$\max_{\lambda\in\Lambda} L(x,\lambda)=Q(x), \tag{8.59}$$

这里 Λ 是对偶问题 (8.44)—(8.45) 的可行域, 即

$$\Lambda=\{\lambda\in\mathbf{R}^m|\lambda_i\geqslant 0,\ i\in I\}. \tag{8.60}$$

对任何 $\lambda\in\Lambda$, 我们有

$$y=A\lambda-g, \tag{8.61}$$

则 (λ,y) 是对偶问题 (8.41)—(8.43) 的可行点, 而且有

$$\min_{x\in\mathbf{R}^n} L(x,\lambda)=b^{\mathrm{T}}\lambda-\frac{1}{2}y^{\mathrm{T}}H^{-1}y=\bar{Q}(\lambda,y). \tag{8.62}$$

设 (x^*,λ^*) 是 (8.28)—(8.32) 的解, 令 $y^*=A\lambda^*-g$, 则知 (λ^*,y^*) 是问题 (8.41)—(8.43) 的可行点, 于是对任何 $x^*\in\mathbf{R}^n$ 和任何 $\lambda\in\Lambda$ 都有

$$\begin{aligned} L(x,\lambda^*)\geqslant&\bar{Q}(\lambda^*,y^*)\\ =&L(x^*,\lambda^*)=Q(x^*)\geqslant L(x^*,\lambda), \end{aligned}\tag{8.63}$$

故知 (x^*,λ^*) 是 $L(x,\lambda)$ 的鞍点. 反之, 如果

$$L(x,\lambda^*)\geqslant L(x^*,\lambda^*)\geqslant L(x^*,\lambda) \tag{8.64}$$

对一切 $x\in X$ 和一切 $\lambda\in\Lambda$ 都成立, 则知

$$-(\lambda-\lambda^*)^{\mathrm{T}}(A^{\mathrm{T}}x^*-b)>0 \tag{8.65}$$

$$\lambda_i\geqslant 0,\quad i\in I \tag{8.66}$$

无解. 利用 Farkas 引理即知 x^* 必是原始问题的可行解. 由 (8.64) 有 $L(x^*,\lambda^*)\geqslant L(x^*,0)$, 故知

$$\sum_{i=1}^{m}\lambda_i^*(a_i^{\mathrm{T}}x^*-b_i)\leqslant 0. \tag{8.67}$$

如果 $\lambda^*\in\Lambda$, 则从 (8.64) 和 (8.67) 可证

$$\begin{aligned} Q(x)=&L(x,\lambda^*)+\sum_{i=1}^{m}\lambda_i^*(a_i^{\mathrm{T}}x^*-b_i)\\ \geqslant&L(x,\lambda^*)\geqslant L(x^*,\lambda^*)\\ =&Q(x^*)-\sum_{i=1}^{m}\lambda_i^*(a_i^{\mathrm{T}}x^*-b_i)\geqslant Q(x^*) \end{aligned}\tag{8.68}$$

对一切 $x \in X$ 都成立. 于是 x^* 是原始问题的极小点. 因此, 我们得到了如下结果.

定理 8.2.3　设 H 正定, 则 $x^* \in X$ 是原始问题的极小点当且仅当存在 $\lambda^* \in \Lambda$ 使得对一切 $x \in X$ 和一切 $\lambda \in \Lambda$ 都有 (8.64) 成立.

8.3　等式约束问题

等式约束的二次规划问题可写成

$$\min_{x\in\mathbf{R}^n} Q(x) = g^{\mathrm{T}}x + \frac{1}{2}x^{\mathrm{T}}Hx, \tag{8.69}$$

$$\text{s.t.}\quad A^{\mathrm{T}}x = b, \tag{8.70}$$

其中 $g \in \mathbf{R}^n, b \in \mathbf{R}^m, A \in \mathbf{R}^{m+n}, H \in \mathbf{R}^{n\times n}$ 且 H 是对称的, 不失一般性, 假定秩 $(A) = m$.

首先, 我们介绍变量消去法. 假定我们已找到变量 x 的一分解 $x = (x_B \quad x_N)^{\mathrm{T}}$. 其中 $x_B \in \mathbf{R}^m$, $x_N \in \mathbf{R}^{n-m}$; 且对应的分解 $A = \begin{bmatrix} A_B \\ A_N \end{bmatrix}$ 使得 A_B 可逆. 利用这一分解, 约束条件 (8.70) 可写成

$$A_B^{\mathrm{T}}x_B + A_N^{\mathrm{T}}x_N = b. \tag{8.71}$$

由于 A_B^{-1} 存在, 故知

$$x_B = (A_B^{-1})^{\mathrm{T}}(b - A_N^{\mathrm{T}}x_N). \tag{8.72}$$

将 (8.72) 代入 (8.69) 就得到 (8.69)—(8.70) 的一个等价形式,

$$\min_{x_N\in\mathbf{R}^{n-m}} \hat{g}_N^{\mathrm{T}}x_N + \frac{1}{2}x_N^{\mathrm{T}}\hat{H}_N x_N, \tag{8.73}$$

其中

$$\hat{g}_N = g_N - A_N A_B^{-1} gB + [H_{NB} - A_N A_B^{-1} H_{BB}](A_B^{-1})^{\mathrm{T}} b, \tag{8.74}$$

$$\hat{H}_N = H_{NN} - H_{NB}(A_B^{-1})^{\mathrm{T}}A_N^{\mathrm{T}} - A_N A_B^{-1} H_{BN} + A_N A_B^{-1} H_{BB}(A_B^{-1})^{\mathrm{T}}A_N^{\mathrm{T}}, \tag{8.75}$$

以及

$$g = \begin{bmatrix} g_B \\ g_N \end{bmatrix}, \tag{8.76}$$

$$H = \begin{bmatrix} H_{BB} & H_{BN} \\ H_{NB} & H_{NN} \end{bmatrix} \tag{8.77}$$

是与 $x = (x_B x_N)^{\mathrm{T}}$ 相应的分解.

如果 $\hat{H}_n$ 正定, 则显然 (8.73) 的解由

$$x_N^* = -\hat{H}_N^{-1}\hat{g}_N \tag{8.78}$$

唯一地给出. 这里, 问题 (8.69)—(8.70) 的解为

$$x^* = \begin{bmatrix} x_B^* \\ x_N^* \end{bmatrix} = \begin{bmatrix} (A_B^{-1})^{\mathrm{T}} b \\ 0 \end{bmatrix} + \begin{bmatrix} (A_B^{-1})^{\mathrm{T}} A_N^{\mathrm{T}} \\ -I \end{bmatrix} \hat{H}_N^{-1}\hat{g}_N. \tag{8.79}$$

设在解 x^* 处的 Lagrange 乘子为 λ^*, 则有

$$g + Hx^* = A\lambda^*. \tag{8.80}$$

从而可知

$$\lambda^* = A_B^{-1}[g_B + H_{BB}x_B^* + H_{BN}x_N^*]. \tag{8.81}$$

$$\min Q(x) = x_1^2 - x_2^2 - x_3^2, \tag{8.82}$$

$$\text{s.t.} \quad x_1 + x_2 + x_3 = 1, \tag{8.83}$$

$$x_2 - x_3 = 1. \tag{8.84}$$

由 (8.84), 可得 x_2 表示为

$$x_2 = x_3 + 1. \tag{8.85}$$

将上式代入 (8.83), 得到

$$x_1 = -2x_3. \tag{8.86}$$

式 (8.85)—(8.86) 实质上就是在变量分解 $x_B = (x_1 x_2), x_N = x_3$ 下所得到的 (8.72). 将 (8.85)—(8.86) 代入 (8.82) 就得到

$$\min_{x_3 \in \mathbf{R}} 4x_3^2 - (x_3 + 1)^2 - x_3^2. \tag{8.87}$$

从上式可得 $x_3 = \dfrac{1}{2}$, 将其代入 (8.85)—(8.86) 就得到了 (8.82)—(8.84) 之解 $\left(-1, \dfrac{3}{2}, \dfrac{1}{2}\right)$. 利用 $g^* = A\lambda^*$ 就可得到

$$\begin{pmatrix} -2 \\ -3 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \lambda_1^* \\ \lambda_2^* \end{pmatrix}. \tag{8.88}$$

从上式可求得 Lagrange 乘子 $\lambda_1^* = -2, \lambda_2^* = -1$.

如果在经过变量消去后的问题 (8.73) 中 $\hat{H}_N$ 是半正定的, 则在

$$(I-\hat{H}_N\hat{H}_N^+)\hat{g}_N=0 \tag{8.89}$$

时, 问题 (8.73) 有界, 且它的解可表示为

$$x_N^*=-\hat{H}_N^+\hat{g}_N+(I-\hat{H}_N^+\hat{H}_N)\tilde{x}. \tag{8.90}$$

其中 $\tilde{x}\in\mathbf{R}^{n-m}$ 是任何向量, H^+ 表示 H 的广义逆矩阵. 在这种情形下, 原问题 (8.69)—(8.70) 的解可用 (8.90) 和 (8.72) 给出. 如果 (8.89) 不成立, 不难发现问题 (8.73) 无下界, 从而原问题 (8.69)—(8.71) 也无下界.

如果 $\hat{H}_N$ 有负特征值, 则很显然 (8.73) 无下界, 故知此时问题 (8.69) 和 (8.70) 不存在有限解.

消去法思想简单明了, 但它的不足之处是 A_B 可能接近一奇异阵, 从而利用 (8.79) 求解 x^* 可能导致数值不稳定.

消去法的一个直接推广是广义消去法. 设 $y_1,\cdots,y_m$ 是域空间 Range(A) 中的一组线性无关的向量, $z_1,\cdots,z_{n-m}$ 是零空间 Null(A^{T}) 中的一组线性无关向量. 记

$$Y=[y_1,\cdots,y_m], \tag{8.91}$$

$$Z=[z_1,\cdots,z_{n-m}]. \tag{8.92}$$

则不难看出, $A^{\mathrm{T}}Y$ 非奇异, $A^{\mathrm{T}}Z=0$. 令

$$x=Y\tilde{x}+Z\hat{x}, \tag{8.93}$$

则从约束条件 (8.70) 即知

$$b=A^{\mathrm{T}}x=A^{\mathrm{T}}Y\tilde{x}. \tag{8.94}$$

所以问题 (8.69)—(8.70) 的可行点可表示为

$$x=Y(A^{\mathrm{T}}Y)^{-1}b+Z\hat{x}, \tag{8.95}$$

其中 $\hat{x}\in\mathbf{R}^{n-m}$ 是自由变量. 将 (8.95) 代入 (8.69) 就得到

$$\min_{\hat{x}\in\mathbf{R}^{n-m}}(g+HY(A^{\mathrm{T}}Y)^{-1}b)^{\mathrm{T}}Z\hat{x}+\frac{1}{2}\hat{x}^{\mathrm{T}}Z^{\mathrm{T}}HZ\hat{x}. \tag{8.96}$$

假定 $Z^{\mathrm{T}}HZ$ 正定, 则从上式可求得解

$$\hat{x}^*=-(Z^{\mathrm{T}}HZ)^{-1}Z^{\mathrm{T}}(g+HY(A^{\mathrm{T}}Y)^{-1}b). \tag{8.97}$$