约束矩阵方程

最佳逼近解的求解方法

杨家稳 著

合肥工业大学出版社

约束矩阵方程最佳逼近解的求解方法

杨家稳　著

合肥工业大学出版社

内容提要

约束矩阵方程的最佳逼近问题是数值代数的重要研究领域之一，它在故障检测、遥感、自动化控制、结构动力学、光谱分析、生物学、电学等领域都有重要的应用。

本书共分五章，分别介绍了约束矩阵方程及其最佳逼近问题的研究现状；本书所用到的基础知识和研究的问题；基于复合最速下降法的迭代算法求约束矩阵方程最佳逼近解；基于正交搜索方向法的迭代算法求约束矩阵方程最佳逼近解；基于余项正交法的迭代算法求约束矩阵方程最佳逼近解。

本书适合于从事数值代数、计算数学等相关领域的大学生、教师和科技工作者阅读。

图书在版编目(CIP)数据

约束矩阵方程最佳逼近解的求解方法/杨家稳著．—合肥：合肥工业大学出版社，2015.7

ISBN 978-7-5650-2353-8

Ⅰ.①约… Ⅱ.①杨… Ⅲ.①矩阵—线性方程—最佳逼近 Ⅳ.①0241.6

中国版本图书馆 CIP 数据核字(2015)第 177352 号

约束矩阵方程最佳逼近解的求解方法

杨家稳 著　　　　责任编辑 魏亮瑜

出　版	合肥工业大学出版社	版　次	2015 年 7 月第 1 版
地　址	合肥市屯溪路 193 号	印　次	2015 年 7 月第 1 次印刷
邮　编	230009	开　本	710 毫米×1010 毫米 1/16
电　话	综合编辑部：0551-62903028	印　张	12.5
	市场营销部：0551-62903198	字　数	229 千字
网　址	www.hfutpress.com.cn	印　刷	安徽联众印刷有限公司
E-mail	hfutpress@163.com	发　行	全国新华书店

ISBN 978-7-5650-2353-8　　　　定价：28.00 元

前　言

约束矩阵方程的最佳逼近问题是数值代数的重要研究领域之一，它在参数识别、特征结构配置、故障检测、遥感、自动化控制、结构动力学、固体力学、光谱分析、生物学、电学、振动理论等领域都有重要的应用。对各类约束矩阵方程最佳逼近问题的研究一直是数值代数研究的热点，越来越受到众多矩阵论和数值分析学者的关注和重视。随着最优化计算方法的发展和计算机的广泛使用，这一领域得到了迅猛发展。特别是近十年来，一大批科技工作者投入了该领域的研究，并且取得了辉煌的成果，一些问题得到了很好的解决，但是有一些问题解决得还不满意，甚至有些问题还没有得到解决。因此，还需要一些科技工作者继续为之努力，发展和完善它的理论体系，拓展它的应用领域。

本书是以2011年安徽高校省级自然科学研究一般项目(Sylvester矩阵方程最佳逼近解的数值方法，项目编号为KJ2011B119)为依托，在作者多年研究的基础上编写而成。本书尝试用一些理论和方法去解决约束矩阵方程的最佳逼近问题。首先，本书希望能够丰富约束矩阵方程最佳逼近理论的内涵；其次，通过对几类约束矩阵方程最佳逼近解的研究，本书给出一些新的理论和算法，希望能够为约束矩阵方程最佳逼近理论的发展和完善贡献一份力量；最后，希望能够为一些科技工作者在解决数值代数等领域的问题时提供一点灵感和帮助。

本书主要探索约束矩阵方程的最佳逼近问题，给出三类迭代算法，不论约束矩阵方程是否有解，这三类算法都能得到约束矩阵方程的最佳逼近解。第一类是基于复合最速下降法的一类迭代算法；第

二类是基于正交搜索方向法的一类迭代算法；第三类是基于余项正交法的一类迭代算法。理论证明这些算法是收敛的，数值例子也表明这些算法是可行的、有效的。

作者诚挚地感谢恩师河海大学理学院的孙合明老师！在本书的编写过程中，作者得到了孙老师的大力支持和悉心指导；感谢师妹胡珊珊和钟青给予的支持和帮助；感谢安徽水利水电职业技术学院的黄建国老师和滁州学院的李晓春老师给予的帮助；感谢合肥工业大学出版社的陆向军主任和魏亮瑜老师为本书出版所做的大量工作；感谢那些默默支持和帮助的亲人和朋友们。

由于作者水平有限，书中如有不妥乃至谬误之处，祈望读者批评指正。

作　者

2015 年 6 月

本书主要符号索引

$\mathbf{R}$	实数集合
int$S$	集合 S 的内点
$\mathbf{R}^n$	n 维实向量集合
$\mathbf{R}^{p\times l}$	$p\times l$ 阶实矩阵集合
$\mathbf{C}^{p\times l}$	$p\times l$ 阶复矩阵集合
real($\mathbf{C}$)	复矩阵的实部
imag($\mathbf{C}$)	复矩阵的虚部
$\boldsymbol{I}$	单位矩阵
$\boldsymbol{A}^{\mathrm{T}}$	矩阵 $\boldsymbol{A}$ 的转置
rank($\boldsymbol{A}$)	矩阵 $\boldsymbol{A}$ 的秩
$R(\boldsymbol{A})$	矩阵 $\boldsymbol{A}$ 的值域
H	实 Hilbert 空间
$\overline{a}$	复数 a 的共轭复数
$\boldsymbol{A}^{\mathrm{H}}$	矩阵 $\boldsymbol{A}$ 的共轭转置
$\mathbf{SR}^{n\times n}$	n 阶实对称矩阵集合
$\mathbf{ASR}^{n\times n}$	n 阶实反对称矩阵集合
$\mathbf{HC}^{n\times n}$	n 阶 Hermite 矩阵集合
$\mathbf{AHC}^{n\times n}$	n 阶反-Hermite 矩阵集合
$\mathbf{GRC}^{n\times n}$	n 阶广义反射矩阵集合
$\mathbf{GRR}^{n\times n}$	n 阶广义实反射矩阵集合
$\mathbf{C}_r^{n\times n}(\boldsymbol{P})$	关于 $\boldsymbol{P}$ 的 n 阶自反矩阵集合
$\mathbf{AC}_r^{n\times n}(\boldsymbol{P})$	关于 $\boldsymbol{P}$ 的 n 阶反自反矩阵集合

（续表）

$\mathrm{vec}\boldsymbol{A}$	矩阵 $\boldsymbol{A}$ 的拉直算子
$\langle \boldsymbol{A}, \boldsymbol{B} \rangle$	矩阵 $\boldsymbol{A}$ 与矩阵 $\boldsymbol{B}$ 的内积
$\\| \boldsymbol{x} \\|_2$	向量 $\boldsymbol{x}$ 的 2-范数或欧式范数
$\\| \boldsymbol{A} \\|$	矩阵 $\boldsymbol{A}$ 的 Frobenius 范数或 F-范数
$\boldsymbol{A} \otimes \boldsymbol{B}$	矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的 Kronecker 积（或直积，或张量积）
$\nabla f(\boldsymbol{x})$	$f(\boldsymbol{x})$在向量 $\boldsymbol{x}$ 处的梯度
$\nabla^2 f(\boldsymbol{x})$	$f(\boldsymbol{x})$在向量 $\boldsymbol{x}$ 处的 Hessian 矩阵
$\nabla f(\boldsymbol{X})$	$f(\boldsymbol{X})$在矩阵 $\boldsymbol{X}$ 处的梯度
$\nabla^2 f(\mathrm{vec}\boldsymbol{X})$	$f(\boldsymbol{X})$在矩阵 $\boldsymbol{X}$ 处的 Hessian 矩阵
$\mathrm{Fix}(T)$	非扩展映射 T 的所有不动点集合
φ'	函数 φ 的 G 导数

目　录

第1章　绪　论

1.1　约束矩阵方程最佳逼近问题的概述

约束矩阵方程最佳逼近问题在参数识别、特征结构配置、故障检测、线性规划和非线性规划、有限元、遥感、自动化控制、结构动力学、固体力学、光谱分析、生物学、电学、振动理论等领域都有重要的应用[1~17]. 由于约束矩阵方程最佳逼近问题的研究有重要意义,因此,对各类约束矩阵方程最佳逼近的研究一直是数值代数研究的热点,已经受到了越来越多矩阵论和数值分析学者们的关注和重视.

约束矩阵方程可以分为相容和不相容两种情况. 在约束矩阵方程相容的情况下,约束矩阵方程的最佳逼近问题就是指在满足特定条件的矩阵集合中求出与已知矩阵最接近的矩阵方程解. 在约束矩阵方程不相容情况下,一般求最小二乘解. 在最小二乘解里求与一个已知矩阵最接近的矩阵,就是矩阵方程的最佳逼近问题. 根据已知逼近的矩阵是零矩阵和非零矩阵,矩阵方程的最佳逼近问题又分为两种情况. 若已知逼近的矩阵为零矩阵,就是求矩阵方程的最小范数最小二乘解. 若已知逼近的矩阵为非零矩阵,就是求矩阵方程的最佳逼近解. 下面列举一类重要的矩阵方程来说明约束矩阵方程及其最佳逼近问题. 自反(反自反)约束矩阵方程 $\boldsymbol{AXB}+\boldsymbol{CX}^{\mathrm{T}}\boldsymbol{D}=\boldsymbol{E}$ 在实验设计和自动化控制中有着很重要的作用. 该约束矩阵方程的最佳逼近问题可描述为:求自反矩阵 $\boldsymbol{X}$ 满足 $\|\boldsymbol{AXB}+\boldsymbol{CX}^{\mathrm{T}}\boldsymbol{D}-\boldsymbol{E}\|=\min$,若满足该条件的 $\boldsymbol{X}$ 的集合为 S_X,在 S_X 中再找出一个 $\boldsymbol{X}^*$ 使得 $\|\boldsymbol{X}^*-\hat{\boldsymbol{X}}\|=\min\|\boldsymbol{X}-\hat{\boldsymbol{X}}\|$ 成立,其中 $\hat{\boldsymbol{X}}$ 是给定矩阵. 该最佳逼近问题就是在约束矩阵方程的最小二乘解中找出一个和已知矩阵 $\hat{\boldsymbol{X}}$ 最接近的矩阵 $\boldsymbol{X}^*$. 该矩阵方程逼近问题出现于实验设计中,矩阵 $\hat{\boldsymbol{X}}$ 来自于试验,但一般不满足矩阵方程 $\boldsymbol{AXB}+\boldsymbol{CX}^{\mathrm{T}}\boldsymbol{D}=\boldsymbol{E}$. 因此,对该类约束矩阵方程 $\boldsymbol{AXB}+\boldsymbol{CX}^{\mathrm{T}}\boldsymbol{D}=\boldsymbol{E}$ 的最佳逼近问题的研究具有现实意义.

对矩阵方程解的约束一般可分为对称(反对称)矩阵、自反(反自反)矩阵、广义自反(反自反)矩阵、P-对称矩阵、双对称矩阵、正交矩阵、反对称正交矩阵、Toeplitz 矩阵等. 解决约束矩阵方程问题的常用方法有 Moore-Penrose(穆尔-彭罗斯)广义逆、奇异值分解(SVD)、广义奇异值分解(GSVD)、商奇异值分解(QSVD)、标准相关分解(CCD)、Cholesky 分解、Shur 分解和迭代法. 使用

Moore-Penrose 广义逆与矩阵分解(SVD、GSVD、QSVD、CCD 等)方法,研究在对称、双对称、中心对称以及对称半正定等约束条件下的矩阵方程问题,得到各类方程可解的充分必要条件以及通解的一般表达式,但是 Moore-Penrose 广义逆和矩阵分解方法对矩阵方程条件的要求比较高,因此极大地限制了这些方法的使用范围.随着自动化控制、航天航空及遥感等科学领域的快速发展,工程计算的量和规模也急剧加大.使用 Moore-Penrose 广义逆或矩阵分解的方法来求解约束矩阵方程问题不仅计算量巨大,而且占用储存空间也很大,因此 Moore-Penrose 广义逆与矩阵分解的方法不适合解决大规模矩阵方程问题.近几年来,用迭代方法解决约束矩阵方程及其最佳逼近问题成为该领域的一个新方向.彭亚新[18] 运用共轭梯度法的思想,提出了一种迭代算法,解决了矩阵方程相容时矩阵方程 $\boldsymbol{AX}=\boldsymbol{B}$,$\boldsymbol{AXB}=\boldsymbol{C}$ 和矩阵方程组 $\begin{cases}\boldsymbol{A}_1\boldsymbol{XB}_1=\boldsymbol{C}_1,\\ \boldsymbol{A}_2\boldsymbol{XB}_2=\boldsymbol{C}_2\end{cases}$ 的对称解、中心对称解、自反解和双对称解等问题.彭卓华[19] 运用了广义共轭梯度法,解决了矩阵方程不相容时矩阵方程 $\boldsymbol{AX}=\boldsymbol{B}$,$\boldsymbol{AXB}=\boldsymbol{C}$ 和矩阵方程组 $\begin{cases}\boldsymbol{A}_1\boldsymbol{XB}_1=\boldsymbol{C}_1,\\ \boldsymbol{A}_2\boldsymbol{XB}_2=\boldsymbol{C}_2\end{cases}$ 的对称解、中心对称解、自反解和对称解等及其最佳逼近问题.雷渊、Mehdi Dehghan 等学者在矩阵方程的约束解、最小二乘解及最佳逼近解方面都做出了有益探索,并取得了一些成果[20~26].

1.2 矩阵方程及其最佳逼近问题的研究现状

1951 年,Bjerhammar[27] 利用广义逆得到矩阵方程 $\boldsymbol{AX}=\boldsymbol{B}$ 一般解存在的充分必要条件和通解的一般表达式.1955 年,Moore-Penrose [28] 利用广义逆得到矩阵方程 $\boldsymbol{AXB}=\boldsymbol{C}$ 一般解存在的充分必要条件和通解的一般表达式.1970 年,Lancaster[29] 利用 Kronecker 积来研究矩阵方程 $\boldsymbol{AXB}=\boldsymbol{C}$ 的一般解与显示解.1984 年,张磊[30] 研究了矩阵方程 $\boldsymbol{AX}=\boldsymbol{B}$ 的正定解,并给出了其可解的条件与通解的一般表达式.1987 年,Henk Don [31] 运用拉直算子研究了矩阵方程 $\boldsymbol{AX}=\boldsymbol{B}$ 的对称解,并给出了其可解的条件及通解的一般表达式.1988 年,Higham[32] 探讨了矩阵方程 $\boldsymbol{AX}=\boldsymbol{B}$ 的最小二乘对称半正定解问题.孙继广[33] 和张磊[34] 分别用不同的方法研究了矩阵方程 $\boldsymbol{AX}=\boldsymbol{B}$ 最小二乘对称解问题.1989 年,Chu[35] 使用矩阵分解方法研究了几个矩阵方程的对称解问题.1990 年,戴华[36] 利用矩阵的奇异值分解方法研究了方程 $\boldsymbol{AX}=\boldsymbol{B}$ 的对称解,并给出了其可解的条件.1992 年,Jameson [37] 研究了矩阵方程 $\boldsymbol{AX}=\boldsymbol{XA}^{\mathrm{T}}$ 和 $\boldsymbol{AX}=\boldsymbol{YB}$ 对称半正定解和对称正定解问题.1993 年,Chang[38] 探讨了矩阵方程组 $\boldsymbol{AX}+\boldsymbol{YA}=\boldsymbol{C}$, $\boldsymbol{AXA}^{\mathrm{T}}+\boldsymbol{BYB}^{\mathrm{T}}=$

C 和 $(A^TXA, B^TYB)=(C,D)$ 的对称解问题. 1993 年，戴华[39] 研究了线性流形上的实对称矩阵最佳逼近问题. 1994 年，戴华[40] 研究了线性流形上的矩阵最佳逼近问题. 1996 年，戴华[41] 利用矩阵的奇异值分解和广义奇异值分解方法研究了方程 $(AX, XB)=(C,D)$ 和 $AXB=C$ 有对称正定解的充分必要条件，并且给出了一般对称正定解的表达式. 1998 年，胡锡炎、张磊与谢冬秀[42] 讨论了在双对称集合中矩阵方程 $AX=B$ 的最小二乘解. 2000 年，廖安平、赵霆雷[43] 及屠文伟[44] 分别研究了矩阵方程 $A^TXA=D$ 的双对称解，并给出了其可解的条件和解的一般表达式. 2001 年，袁永新[45] 利用广义奇异值分解方法研究了矩阵方程 $AXB+CYD=F$ 的最小范数解. 2002 年，Konstantinos Slavakis[46] 使用最速下降法研究了对称正定 Toeplitz 解及其最佳逼近问题. 2002 年，廖安平和白中治[47] 利用标准奇异值分解方法研究了矩阵方程 $A^TXA=D$ 的双对称最小二乘解问题. 2002 年，彭振赟[48] 研究了矩阵方程 $A^TXA=B$ 的中心对称解及其最佳逼近问题. 2003 年，谢冬秀[49] 利用闭凸锥的理论讨论了 $A^TXA=B$ 最小二乘半正定解及其最佳逼近解问题. 2003 年，彭亚新、胡锡炎和张磊[50] 研究了矩阵方程 $A^TXA=B$ 在线性流形上的对称正交对称解及其最佳逼近，并给出可解的条件和解的一般表达式. 2003 年，邓远北[51] 研究了矩阵方程 $A^TXA=B$ 在线性流形上的对称、反对称解、对称半正定解及其最佳逼近问题. 2004 年，彭亚新[52] 和彭向阳[53] 分别使用矩阵的广义奇异值分解方法得到了 $A^TXA=B$ 矩阵方程有反对称正交反对称解的充分必要条件及其通解的表达式，并且给出了矩阵方程的解集合中与给定矩阵的最佳逼近解. 2004 年，彭亚新[18] 利用共轭梯度法的思想，提出一种迭代算法，解决了矩阵方程相容时矩阵方程 $AX=B$，$AXB=C$ 和矩阵方程组 $\begin{cases}A_1XB_1=C_1,\\A_2XB_2=C_2\end{cases}$ 的对称解、中心对称解、自反解和双对称解等问题. 2004 年，孟纯军[54] 研究了在闭凸锥上 $AX=B$ 的最小二乘非负解和最小二乘半正定解的问题. 2005 年，廖平安[55] 利用奇异值分解方法给出了矩阵方程 $AXA^T+BYB^T=C$ 的对称与反对称最小范数最小二乘解. 2005 年，Peng、Hu 和 Zhang[56] 使用迭代方法研究了 $AXB=C$ 的对称解及其最佳逼近问题. 2005 年，Meng[57] 研究了 $AX=B$ 的斜对称正交解及其最佳逼近问题. 2006 年，龚丽莎[58] 讨论了在子矩阵约束下矩阵方程 $AX=B$ 的最小二乘解. 2006 年，袁永新[59] 利用广义奇异值分解方法给出了 $AXA^T+BYB^T+AZB^T=D$ 和 $AXA^T+AZB^T+BZ^TA^T=D$ 的最小范数对称解. 2007 年，Yuan 和 Liao[60] 用迭代算法给出矩阵方程 $AXB=C$ 在不相容情况下的最小二乘对称解. 2007 年，Wang[61] 用迭代方法研究了 $AXB+CX^TD=E$ 的最小二乘解及其最佳逼近解. 2007 年，Piao[62] 利用 Moore-Penrose 广义逆得到矩阵方程 $AX+X^TC=B$ 的解. 2007 年，袁仕芳[63] 给出矩阵方程 $AXB+CYD=E$ 的对称最小范数最小二乘解. 2007 年，Peng[64] 利用共轭梯度法思想研究了

矩阵方程 $\boldsymbol{A}_1\boldsymbol{X}_1\boldsymbol{B}_1+\boldsymbol{A}_2\boldsymbol{X}_2\boldsymbol{B}_2+\cdots+\boldsymbol{A}_l\boldsymbol{X}_l\boldsymbol{B}_l=\boldsymbol{C}$ 的双对称解及其最佳逼近问题. 2008 年,盛兴平[65] 用迭代方法研究了 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{X}\boldsymbol{B}+\boldsymbol{B}^{\mathrm{T}}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{D}$ 的最小范数最小二乘解. 2008 年,刘大瑾[66] 用迭代算法讨论了 $\boldsymbol{AXB}+\boldsymbol{CXD}=\boldsymbol{F}$ 的中心对称解及其最佳逼近解. 2008 年,Mehdi Dehghan[21] 用迭代方法研究了$(\boldsymbol{AY}-\boldsymbol{ZB},\boldsymbol{CY}-\boldsymbol{ZD})=(\boldsymbol{E},\boldsymbol{F})$ 的自反解和反自反解及其最佳逼近问题. 2008 年,Huang[67] 用迭代方法探讨了矩阵方程 $\boldsymbol{AXB}=\boldsymbol{C}$ 的斜对称解及其最佳逼近问题. 2009 年,Mehdi Dehghan[22] 用迭代方法给出了 $\boldsymbol{A}_1\boldsymbol{X}_1\boldsymbol{B}_1+\boldsymbol{A}_2\boldsymbol{X}_2\boldsymbol{B}_2=\boldsymbol{C}$ 在相容情况下的自反最小范数解和反自反最小范数解. 2009 年,Li[68] 研究了矩阵方程 $\boldsymbol{AXB}+\boldsymbol{CYD}=\boldsymbol{E}$ 的子矩阵约束解问题. 2009 年,Zhang[69] 用迭代方法求相容矩阵方程 $\boldsymbol{AX}=\boldsymbol{B}$ 的$(\boldsymbol{P},\boldsymbol{Q})$ 广义自反最小范数解和反自反最小范数解. 2010 年,Mehdi Dehghan[23] 用迭代方法探讨了矩阵方程组的广义双对称最小范数解. 2010 年,Peng[70] 用迭代算法研究了 $\boldsymbol{AXB}=\boldsymbol{C}$ 的约束最小范数解. 2011 年,Li[71] 讨论了矩阵方程 $\boldsymbol{AXB}=\boldsymbol{C}$ 在特殊子矩阵约束下的中心对称最小范数解. 2011 年,Maziar Dehghan[72] 用两种迭代方法求解矩阵方程的双对称和斜对称最小范数解. 2012 年,Peng[73] 用迭代方法探讨在 $\boldsymbol{CXD}\geqslant\boldsymbol{E}$ 约束下矩阵方程 $\boldsymbol{AX}=\boldsymbol{B}$ 的解. 2012 年,孙合明[73] 利用复合最速下降法研究矩阵方程 $\boldsymbol{AXB}+\boldsymbol{CYD}=\boldsymbol{E}$ 的自反解及其最佳逼近问题. 2013 年,杨家稳[74] 利用最速下降法研究矩阵方程 $\boldsymbol{AXB}+\boldsymbol{CX}^{\mathrm{T}}\boldsymbol{D}=\boldsymbol{E}$ 的自反解及其最佳逼近问题. 2013 年,Peng[75] 利用迭代方法求子矩阵约束下矩阵方程组的最小二乘解. 2013 年,周硕[76] 利用奇异值分解方法研究了中心主子矩阵约束下矩阵方程的中心对称解. 2013 年,Liang[77] 研究复矩阵 $\boldsymbol{AXB}=\boldsymbol{C}$ 的广义自反和反自反解,并给出可解条件和通解的一般表达式. 2014 年,杨家稳[78] 用最速下降法研究了复矩阵方程 $\boldsymbol{A}_1\boldsymbol{Z}+\boldsymbol{ZB}_1=\boldsymbol{C}_1$ 的广义自反最佳逼近问题. 2014 年,赵冰艳[79] 利用矩阵的正交三角分解方法研究矩阵方程组正交$(\boldsymbol{P},\boldsymbol{Q})$-对称解,并给出有正交$(\boldsymbol{P},\boldsymbol{Q})$-对称解的充分必要条件及通解的一般表达式. 2014 年,梁艳芳[80] 使用 Moore-Penrose 广义逆探讨矩阵方程 $\boldsymbol{AXB}+\boldsymbol{CYD}=\boldsymbol{E}$ 的双中心最小二乘问题,得到双中心最小范数最小二乘解和对称双中心最小范数最小二乘解的一般表达式. 2014 年,刘莉[81] 用迭代算法研究了矩阵 $\boldsymbol{AXB}+\boldsymbol{CYD}=\boldsymbol{E}$ 的双对称最小二乘解及其最佳逼近问题. 2015 年,杨家稳[82] 用正交方向法研究了矩阵方程 $\boldsymbol{AXB}+\boldsymbol{CYD}=\boldsymbol{E}$ 的自反解及其最佳逼近问题.

矩阵方程及其最佳逼近解的研究成果很多,特别是近几年,矩阵方程最佳逼近解的研究成果很丰硕.

第 2 章　基础知识与研究的问题

本章将给出本书中所用到的一些基本概念、基本定理和研究的问题.

2.1　基本概念

在本书中,$\mathbf{R}$ 为实数集合, int$S$ 表示集合 S 的内点,$\mathbf{R}^n$ 为 n 维实向量集合,$\mathbf{R}^{p\times l}$ 为 $p\times l$ 阶实矩阵集合,$\mathbf{C}^{p\times l}$ 为 $p\times l$ 阶复矩阵集合,real($\mathbf{C}$) 为复矩阵的实部,imag($\mathbf{C}$) 为复矩阵的虚部,$\boldsymbol{I}$ 为单位矩阵,$\boldsymbol{A}^{\mathrm{T}}$ 表示矩阵 $\boldsymbol{A}$ 的转置,rank($\boldsymbol{A}$) 表示矩阵 $\boldsymbol{A}$ 的秩,$R(\boldsymbol{A})$ 表示矩阵 $\boldsymbol{A}$ 的值域,H 为实 Hilbert 空间,$\bar{a}$ 表示复数 a 的共轭复数,$\boldsymbol{A}^{\mathrm{H}}$ 表示矩阵 $\boldsymbol{A}$ 的共轭转置.

定义 2.1　设 $\boldsymbol{A}\in\mathbf{R}^{n\times n}$,若 $\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}$,则称 $\boldsymbol{A}$ 为实对称矩阵(real symmetric matrix). 记实对称矩阵全体组成的集合为 $\mathbf{SR}^{n\times n}$, 即 $\mathbf{SR}^{n\times n}=\{\boldsymbol{X}\in\mathbf{R}^{n\times n}\mid\boldsymbol{X}=\boldsymbol{X}^{\mathrm{T}}\}$.

定义 2.2　设 $\boldsymbol{A}\in\mathbf{R}^{n\times n}$,若 $\boldsymbol{A}^{\mathrm{T}}=-\boldsymbol{A}$,则称 $\boldsymbol{A}$ 为实反对称矩阵(real anti-symmetric matrix). 记实反对称矩阵全体组成的集合为 $\mathbf{ASR}^{n\times n}$,即 $\mathbf{ASR}^{n\times n}=\{\boldsymbol{X}\in\mathbf{R}^{n\times n}\mid\boldsymbol{X}=-\boldsymbol{X}^{\mathrm{T}}\}$.

定义 2.3　设 $\boldsymbol{A}=(a_{ij})_{n\times n}\in\mathbf{C}^{n\times n}$,若 $\boldsymbol{A}^{\mathrm{H}}=\boldsymbol{A}$,则称 $\boldsymbol{A}$ 为 Hermite 矩阵 (Hermiterian matrices). 记 Hermite 矩阵全体组成的集合为 $\mathbf{HC}^{n\times n}$,即 $\mathbf{HC}^{n\times n}=\{\boldsymbol{X}\in\mathbf{C}^{n\times n}\mid\boldsymbol{X}=\boldsymbol{X}^{\mathrm{H}}\}$.

定义 2.4　设 $\boldsymbol{A}=(a_{ij})_{n\times n}\in\mathbf{C}^{n\times n}$,若 $\boldsymbol{A}^{\mathrm{H}}=-\boldsymbol{A}$,则称 $\boldsymbol{A}$ 为反-Hermite 矩阵 (anti-Hermiterian matrices). 记反- Hermite 矩阵全体组成的集合为 $\mathbf{AHC}^{n\times n}$,即 $\mathbf{AHC}^{n\times n}=\{\boldsymbol{X}\in\mathbf{C}^{n\times n}\mid\boldsymbol{X}=-\boldsymbol{X}^{\mathrm{H}}\}$.

定义 2.5　设 $\boldsymbol{P}\in\mathbf{C}^{n\times n}$,若 $\boldsymbol{P}^2=\boldsymbol{I}$ 且 $\boldsymbol{P}^{\mathrm{H}}=\boldsymbol{P}$,则称矩阵 $\boldsymbol{P}$ 是广义反射矩阵 (generalized reflection matrix). 记广义反射矩阵全体组成的集合为 $\mathbf{GRC}^{n\times n}$,即 $\mathbf{GRC}^{n\times n}=\{\boldsymbol{X}\in\mathbf{C}^{n\times n}\mid\boldsymbol{X}^2=\boldsymbol{I}\text{ 且 }\boldsymbol{X}^{\mathrm{H}}=\boldsymbol{X}\}$.

定义 2.6　设 $\boldsymbol{P}\in\mathbf{R}^{n\times n}$,若 $\boldsymbol{P}^2=\boldsymbol{I}$ 且 $\boldsymbol{P}^{\mathrm{T}}=\boldsymbol{P}$,则称矩阵 $\boldsymbol{P}$ 是广义实反射矩阵 (generalized real reflection matrix). 记广义实反射矩阵全体组成的集合为 $\mathbf{GRR}^{n\times n}$,即 $\mathbf{GRR}^{n\times n}=\{\boldsymbol{X}\in\mathbf{R}^{n\times n}\mid\boldsymbol{X}^2=\boldsymbol{I}\text{ 且 }\boldsymbol{X}^{\mathrm{T}}=\boldsymbol{X}\}$.

定义 2.7　设 $\boldsymbol{P}\in\mathbf{GRC}^{n\times n}$,$\boldsymbol{A}\in\mathbf{C}^{n\times n}$,若 $\boldsymbol{PAP}=\boldsymbol{A}$ 或者 $\boldsymbol{PAP}=-\boldsymbol{A}$,则称矩阵 $\boldsymbol{A}$ 为关于 $\boldsymbol{P}$ 的自反矩阵 (reflexive matrices) 或者关于 $\boldsymbol{P}$ 的反自反矩阵

(anti-reflexive matrices). 分别记关于 $\boldsymbol{P}$ 的自反矩阵或者反自反矩阵全体组成的集合为 $\mathbf{C}_r^{n\times n}(\boldsymbol{P})$ 或者 $\mathbf{AC}_r^{n\times n}(\boldsymbol{P})$，即

$$\mathbf{C}_r^{n\times n}(\boldsymbol{P})=\{\boldsymbol{X}\in\mathbf{C}^{n\times n}\mid \boldsymbol{PXP}=\boldsymbol{X}\},\ \mathbf{C}_r^{n\times n}(\boldsymbol{P})=\{\boldsymbol{X}\in\mathbf{C}^{n\times n}\mid \boldsymbol{PXP}=-\boldsymbol{X}\}.$$

定义 2.8 若矩阵 $\boldsymbol{A}=(a_{ij})_{m\times n}$，称$(a_{11},\cdots,a_{m1},a_{12},\cdots a_{1n},\cdots,a_{mn})^{\mathrm{T}}$ 为矩阵 $\boldsymbol{A}$ 的拉直算子，记为 $\mathrm{vec}\boldsymbol{A}$，即 $\mathrm{vec}\boldsymbol{A}=(a_{11},\cdots,a_{m1},a_{12},\cdots a_{1n},\cdots,a_{mn})^{\mathrm{T}}$.

定义 2.9 设矩阵 $\boldsymbol{A}=(a_{ij})\in\mathbf{C}^{m\times n}$，$\boldsymbol{B}=(b_{ij})\in\mathbf{C}^{m\times n}$，则称 $\sum\limits_{i=1}^{m}\sum\limits_{j=1}^{n}a_{ij}\bar{b}_{ij}$ 为矩阵 $\boldsymbol{A}$ 与矩阵 $\boldsymbol{B}$ 的内积，记作$\langle\boldsymbol{A},\boldsymbol{B}\rangle$，即$\langle\boldsymbol{A},\boldsymbol{B}\rangle=\sum\limits_{i=1}^{m}\sum\limits_{j=1}^{n}a_{ij}\bar{b}_{ij}$. 当矩阵 $\boldsymbol{A}=(a_{ij})\in\mathbf{R}^{m\times n}$，$\boldsymbol{B}=(b_{ij})\in\mathbf{R}^{m\times n}$ 时，$\langle\boldsymbol{A},\boldsymbol{B}\rangle=\sum\limits_{i=1}^{m}\sum\limits_{j=1}^{n}a_{ij}b_{ij}$.

定义 2.10 设向量 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}\in\mathbf{R}^n$，称$\sqrt{\langle\boldsymbol{x},\boldsymbol{x}\rangle}$ 为向量 $\boldsymbol{x}$ 的 2-范数或欧式范数，记作 $\|\boldsymbol{x}\|_2$，即 $\|\boldsymbol{x}\|_2=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=\sqrt{x_1^2+x_2^2+\cdots+x_n^2}$.

定义 2.11 设矩阵 $\boldsymbol{A}=(a_{ij})\in\mathbf{C}^{m\times n}$，称$\sqrt{\langle\boldsymbol{A},\boldsymbol{A}\rangle}$ 为矩阵 $\boldsymbol{A}$ 的 Frobenius 范数或 F-范数，记作 $\|\boldsymbol{A}\|$，即 $\|\boldsymbol{A}\|=\sqrt{\langle\boldsymbol{A},\boldsymbol{A}\rangle}$.

定义 2.12 设非零矩阵 $\boldsymbol{A}\in\mathbf{C}^{m\times n}$ 和 $\boldsymbol{B}\in\mathbf{C}^{m\times n}$，若 $\boldsymbol{A}\neq\boldsymbol{B}$ 且$\langle\boldsymbol{A},\boldsymbol{B}\rangle=0$，称矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 正交.

定义 2.13 设矩阵 $\boldsymbol{A}=(a_{ij})\in\mathbf{C}^{m\times n}$，$\boldsymbol{B}=(b_{ij})\in\mathbf{C}^{m\times n}$，则称如下的分块矩阵

$$\boldsymbol{A}\otimes\boldsymbol{B}=\begin{pmatrix}a_{11}\boldsymbol{B} & a_{12}\boldsymbol{B} & \cdots & a_{1l}\boldsymbol{B}\\ a_{21}\boldsymbol{B} & a_{22}\boldsymbol{B} & \cdots & a_{2l}\boldsymbol{B}\\ \vdots & \vdots & & \vdots\\ a_{p1}\boldsymbol{B} & a_{p2}\boldsymbol{B} & \cdots & a_{pl}\boldsymbol{B}\end{pmatrix}$$

为矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的 Kronecker 积(或直积，或张量积).

注:矩阵乘积与Kronecker积的混合运算顺序是先运算矩阵乘积，然后再运算 Kronecker 积. 例如 $\boldsymbol{AB}\otimes\boldsymbol{CD}$ 的运算顺序是先运算 $\boldsymbol{AB}$ 和 $\boldsymbol{CD}$，然后再运算 Kronecker 积，即 $\boldsymbol{AB}\otimes\boldsymbol{CD}=(\boldsymbol{AB})\otimes(\boldsymbol{CD})$.

定义 2.14 设 S 是 H 的一个子集合，对任意 $x_1,x_2\in S$，及 $\lambda\in[0,1]$，如果 $\lambda x_1+(1-\lambda)x_2\in S$，则称 S 是 H 的一个凸集. 若 S 中的任意序列$\{x_i\}$收敛于 x_0，且 $x_0\in S$，则称 S 为闭集. 既为闭集又为凸集的集合称为闭凸集.

定义 2.15 设 $S\subset H$ 是非空凸集，$\lambda\in[0,1]$，$f(x)$ 是定义在 S 上的函数，如果对任意 $x_1,x_2\in S$，有 $f[\lambda x_1+(1-\lambda)x_2]\leqslant\lambda f(x_1)+(1-\lambda)f(x_2)$，则称函数 $f(x)$ 是 S 上的凸函数.

定义 2.16 设 $f:S\subset\mathbf{R}^n\to\mathbf{R}$，$\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$，以 $f(\boldsymbol{x})$ 的 n 个偏导数

为分量的 n 维向量称为 $f(\boldsymbol{x})$ 在向量 $\boldsymbol{x}$ 处的梯度，记为

$$\nabla f(\boldsymbol{x})=\left(\frac{\partial f(\boldsymbol{x})}{\partial x_1},\frac{\partial f(\boldsymbol{x})}{\partial x_2},\cdots,\frac{\partial f(\boldsymbol{x})}{\partial x_n}\right)^{\mathrm{T}}.$$

定义 2.17　设 $f:S\subset\mathbf{R}^n\to\mathbf{R},\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$，若 $f(\boldsymbol{x})$ 二阶可导，则称 $f(\boldsymbol{x})$ 的二阶导数为 Hessian 矩阵，记为 $\nabla^2 f(\boldsymbol{x})$，即

$$\nabla^2 f(\boldsymbol{x})=\begin{pmatrix}\frac{\partial^2 f(\boldsymbol{x})}{\partial x_1{}^2} & \frac{\partial^2 f(\boldsymbol{x})}{\partial x_2\partial x_1} & \cdots & \frac{\partial^2 f(\boldsymbol{x})}{\partial x_n\partial x_1}\\ \frac{\partial^2 f(\boldsymbol{x})}{\partial x_1\partial x_2} & \frac{\partial^2 f(\boldsymbol{x})}{\partial x_2{}^2} & \cdots & \frac{\partial^2 f(\boldsymbol{x})}{\partial x_n\partial x_2}\\ \vdots & \vdots & & \vdots\\ \frac{\partial^2 f(\boldsymbol{x})}{\partial x_1\partial x_n} & \frac{\partial^2 f(\boldsymbol{x})}{\partial x_2\partial x_n} & \cdots & \frac{\partial^2 f(\boldsymbol{x})}{\partial x_n{}^2}\end{pmatrix}.$$

定义 2.18　设 $f:S\subset\mathbf{R}^{m\times n}\to\mathbf{R}$，矩阵 $\boldsymbol{X}=\begin{pmatrix}x_{11} & x_{12} & \cdots & x_{1n}\\ x_{21} & x_{22} & \cdots & x_{2n}\\ \vdots & \vdots & & \vdots\\ x_{m1} & x_{m2} & \cdots & x_{mn}\end{pmatrix}$，则称

$$\begin{pmatrix}\frac{\partial f(\boldsymbol{X})}{\partial x_{11}} & \frac{\partial f(\boldsymbol{X})}{\partial x_{12}} & \cdots & \frac{\partial f(\boldsymbol{X})}{\partial x_{1n}}\\ \frac{\partial f(\boldsymbol{X})}{\partial x_{21}} & \frac{\partial f(\boldsymbol{X})}{\partial x_{22}} & \cdots & \frac{\partial f(\boldsymbol{X})}{\partial x_{2n}}\\ \vdots & \vdots & & \vdots\\ \frac{\partial f(\boldsymbol{X})}{\partial x_{m1}} & \frac{\partial f(\boldsymbol{X})}{\partial x_{m2}} & \cdots & \frac{\partial f(\boldsymbol{X})}{\partial x_{mn}}\end{pmatrix}$$

为 $f(\boldsymbol{X})$ 在矩阵 $\boldsymbol{X}$ 处的梯度，记为 $\nabla f(\boldsymbol{X})$.

定义 2.19　设 $f:S\subset\mathbf{R}^{m\times n}\to\mathbf{R}$，矩阵 $\boldsymbol{X}=\begin{pmatrix}x_{11} & x_{12} & \cdots & x_{1n}\\ x_{21} & x_{22} & \cdots & x_{2n}\\ \vdots & \vdots & & \vdots\\ x_{m1} & x_{m2} & \cdots & x_{mn}\end{pmatrix}$，则称 $\nabla^2 f(\mathrm{vec}\boldsymbol{X})$ 为 $f(\boldsymbol{X})$ 的 Hessian 矩阵，记为 $\nabla^2 f(\boldsymbol{X})$.

定义 2.20　设 $T:H\to H$ 是定义在 $S\subset H$ 上的一映射，若存在 $\kappa>0$，对任意 $x,y\in S$，有 $\|T(x)-T(y)\|\leqslant\kappa\|x-y\|$，则称 T 为 κ-Lipschitzian. 特别地，对任意 $x,y\in H$，有 $\|T(x)-T(y)\|\leqslant\|x-y\|$，则称 $T:H\to H$ 为非扩展映射. 若对任意的 $x\in H$，有 $T(x)=x$，则称为 T 的所有不动点集合，记

为 $\mathrm{Fix}(T)$.

定义 2.21 给定一个集合 $S\subset H$,$T:H\to H$ 是定义在 S 上的一个映射,若存在 $\eta>0$,对任意 $x,y\in S$,有 $\langle T(x)-T(y),x-y\rangle\geqslant\eta\|x-y\|^2$,则称 T 是强单调的.

定义 2.22 设 U 为 H 的一个开子集,若对任意的 $u\in U$,$h\in H$,存在 $a(u)\in H$,使得 $\lim\limits_{\delta\to 0}\dfrac{\varphi(u+\delta h)-\varphi(u)}{\delta}=\langle a(u),h\rangle$,则称函数 $\varphi:H\to\mathbf{R}\cup\{\infty\}$ 在 U 上是 G 可微的,$\varphi':U\to H:u\mapsto a(u)$ 称为在 U 上的 G 导数.

定义 2.23 设 S 是 H 的一个子集合,对任意给定的 $x\in H$,如果存在 $x_0\in S$,使得 $\|x-x_0\|=\inf\limits_{z\in S}\|x-z\|$,则称 x_0 为 S 中 x 的一个最佳逼近元.

定义 2.24 称满足 $\|\boldsymbol{B}-\boldsymbol{A}\boldsymbol{X}^*\|=\min\|\boldsymbol{B}-\boldsymbol{A}\boldsymbol{X}\|$ 的 $\boldsymbol{X}^*$ 为矩阵方程 $\boldsymbol{A}\boldsymbol{X}=\boldsymbol{B}$ 的最小二乘解.

定义 2.25 若 S 是矩阵方程的解集,$\boldsymbol{X}^*\in S$,且满足 $\|\boldsymbol{X}^*\|=\min\limits_{\boldsymbol{X}\in S}\|\boldsymbol{X}\|$,称 $\boldsymbol{X}^*$ 为矩阵方程的最小范数解或极小范数解.

一般说来,不相容矩阵方程的最小二乘解是不唯一的,但在最小二乘解的集合中,具有最小范数的解是唯一的,此解称为最小范数最小二乘解,或者称为极小范数最小二乘解,或者称为对零矩阵的最佳逼近解.

2.2 基本定理

定理 2.2.1[83] 设矩阵 $\boldsymbol{A}\in\mathbf{C}^{m\times n}$,$\boldsymbol{B}\in\mathbf{C}^{p\times q}$,$\boldsymbol{C}\in\mathbf{C}^{r\times s}$,$\boldsymbol{D}\in\mathbf{C}^{p\times q}$,则

(1) 纯量积:$\lambda(\boldsymbol{A}\otimes\boldsymbol{B})=(\lambda\boldsymbol{A})\otimes\boldsymbol{B}=\boldsymbol{A}\otimes(\lambda\boldsymbol{B})$,$\forall\lambda\in\mathbf{R}$;

(2) 分配律:$(\boldsymbol{A}+\boldsymbol{B})\otimes\boldsymbol{C}=(\boldsymbol{A}\otimes\boldsymbol{C})+(\boldsymbol{B}\otimes\boldsymbol{C})$,$\boldsymbol{C}\otimes(\boldsymbol{A}+\boldsymbol{B})=(\boldsymbol{C}\otimes\boldsymbol{A})+(\boldsymbol{C}\otimes\boldsymbol{B})$;

(3) 转置及共轭:$(\boldsymbol{A}\otimes\boldsymbol{B})^{\mathrm{T}}=\boldsymbol{A}^{\mathrm{T}}\otimes\boldsymbol{B}^{\mathrm{T}}$,$(\boldsymbol{A}\otimes\boldsymbol{B})^{\mathrm{H}}=\boldsymbol{A}^{\mathrm{H}}\otimes\boldsymbol{B}^{\mathrm{H}}$;

(4) 混合积:当 $n=r$,$q=k$ 时,有

$$(\boldsymbol{A}\otimes\boldsymbol{B})(\boldsymbol{C}\otimes\boldsymbol{D})=(\boldsymbol{AC})\otimes(\boldsymbol{BD})=\boldsymbol{AC}\otimes\boldsymbol{BD}.$$

定理 2.2.2 矩阵 $\boldsymbol{A}\in\mathbf{R}^{m\times n}$,$\boldsymbol{X}\in\mathbf{R}^{n\times s}$,$\boldsymbol{B}\in\mathbf{R}^{s\times t}$,$\boldsymbol{C}\in\mathbf{R}^{m\times t}$,则

$$\langle\boldsymbol{AXB},\boldsymbol{C}\rangle=\langle\boldsymbol{X},\boldsymbol{A}^{\mathrm{T}}\boldsymbol{CB}^{\mathrm{T}}\rangle.$$

证明 $\langle\boldsymbol{AXB},\boldsymbol{C}\rangle=\langle\mathrm{vec}(\boldsymbol{AXB}),\mathrm{vec}\boldsymbol{C}\rangle=\langle(\boldsymbol{B}^{\mathrm{T}}\otimes\boldsymbol{A})\mathrm{vec}\boldsymbol{X},\mathrm{vec}\boldsymbol{C}\rangle$

$$=[(\boldsymbol{B}^{\mathrm{T}}\otimes\boldsymbol{A})\mathrm{vec}\boldsymbol{X}]^{\mathrm{T}}\mathrm{vec}\boldsymbol{C}=(\mathrm{vec}\boldsymbol{X})^{\mathrm{T}}(\boldsymbol{B}\otimes\boldsymbol{A}^{\mathrm{T}})\mathrm{vec}\boldsymbol{C}$$