国家科学技术学术著作出版基金
NFAPST

系统与控制理论中的线性代数

（下 册）（第二版）

黄 琳 编著

科学出版社

国家科学技术学术著作出版基金资助出版

系统与控制理论中的线性代数

(下册)

(第二版)

黄　琳　编著

科学出版社

北　京

内 容 简 介

本书为《系统与控制理论中的线性代数》的第二版，保留了原书的基本理论，删除了不必要的内容，增加了近三十年来出现的新的重要理论。书中一些内容是作者长期研究的结果。本书分上下两册，共十三章。上册为基础理论，前四章概述与深化了线性代数的基本理论，后四章为几个重要的特殊理论。下册为应用部分，分别是数值代数的基础，关于稳定性和系统描述与设计涉及的内容，以及一些特殊的矩阵类、S 过程和线性矩阵不等式。各章均附有习题。

本书可供从事应用数学、系统工程与系统理论、控制理论与控制工程、力学和其他应用学科的教学与科研人员参考，亦可作为研究生教材。

图书在版编目(CIP)数据

系统与控制理论中的线性代数.下册/黄琳编著. —2 版. —北京：科学出版社，2018.2

ISBN 978-7-03-056399-6

Ⅰ. ①系… Ⅱ. ①黄… Ⅲ. ①控制论-线性代数计算法 Ⅳ. ①O241.6

中国版本图书馆 CIP 数据核字(2018) 第 013475 号

责任编辑：魏英杰／责任校对：桂伟利
责任印制：师艳茹／封面设计：陈 敬

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

北京通州皇家印刷厂 印刷
科学出版社发行 各地新华书店经销
*
1984 年 1 月第 一 版 开本：720 × 1000 1/16
2018 年 2 月第 二 版 印张：26
2018 年 2 月第四次印刷 字数：525 000

定价：180.00 元

(如有印装质量问题，我社负责调换)

第二版序

三十多年前, 科学出版社出版了《系统与控制理论中的线性代数》. 该书自出版以后, 便得到业界的肯定与抬爱, 还有一些成名学者说是这本书帮助他们进入了现代控制理论的大门, 因此问我 “你当初怎么会想到要出这样一本书的?” 对这个问题简单地回答就是: “这是那个时代需求决定的.” 具体来说, 决定写这本书的动机主要来自以下几点:

(1) 在我国从 20 世纪 50 年代后期至 “文化大革命” 结束的这段时期, 恰好是国际上控制理论、系统工程和计算力学大发展的时期. 三方面的代表性事件分别是: R.E. Kalman 在 1960 年 IFAC 首次会议上提出的关于系统可控可观测的基本概念和随后发展起来的以较多数学为主要研究手段的现代控制理论, 特别是线性系统的状态空间理论; H.W. Kuhn 和 A.W. Tucker 1951 年在 UC Berkeley 举行的 Symposium 上提交的报告 “Nonlinear programming” 中提出的 Kuhn-Tucker 条件为标志和随后开展起来的凸优化研究; 最早由 R. Courant 在 1943 年提出的求解偏微分方程的有限元思想, 后来我国冯康在 60 年代结合水利工程发展出有限元方法, 到了这一时期由于计算机技术的巨大进步而发展成为一门新型的学科——计算力学. 这些典型事件有些是在此前发生的, 但经过近 20 年的发酵与演进, 到了 80 年代初都发展到了相当的规模, 而这段时间刚好是我们痛失发展良机的时候, 使得我们原本并不先进的科技与国际水平的差距进一步拉大了.

(2) 就以上三方面的发展来说, 线性情形总是发展的基础, 将大量的工程和物理系统的问题在相当广的范围里看作是线性问题是合理的. 因此, 在以上三方面要能赶上世界发展的节奏, 当务之急是首先解决在线性情形下的差距, 没有这个好的基础其他一切都很难谈起. 如果说 60 年代以前的控制或力学范畴的书和论文, 出现线性代数的描述或用矩阵工具还并不多见, 而到了 70 年代后期情况已发生了根本的变化. 可以想见, 要想用数学工具解决上述三方面, 特别是控制与系统的理论或实际问题, 线性代数当是首选工具。

(3) 所有非线性问题的解决, 实际上都离不开线性情形的方法与理论, 这不仅是由于用线性情形的累积在很多情况下可以去逼近非线性情形, 而且有些非线性理论实际是按线性理论的框架建立的. 例如, 20 世纪 80 年代前后发展起来的以微分几何为主要研究方法的非线性控制理论实际上是微分几何与线性控制的结合. 有人建议如果把微分几何和线性代数从概念上作一比较并建立起对应关系, 则有线性系统知识的人抓住这种对应就能很自然地理解了这种非线性控制理论, 这也解释了为

什么这种非线性理论会吸引那么多人研究, 得到那么多结果却基本上不适合解决系统中本质非线性, 如自振、混沌等问题. 即使对系统本质非线性问题的研究, 无论是理论推演与论证, 还是实际计算, 离开了线性的工具也必然一筹莫展.

(4) 计算机的快速发展使其计算能力日新月异, 50 年代前那种认为 10 阶以上矩阵特征值的计算几乎是无法进行下去的断言已一去不复返. 计算方法与线性代数的结合出现了一个新的数值线性代数的学科. 由于航空航天的迅速发展和工业过程控制的进步都使得面临的系统与控制问题日益复杂, 其研究已不能仅靠由频率特性派生出的简单列线图加以解决. 控制科学、优化技术和计算力学本质上当然都是技术科学, 衡量其是否有价值首先应该是有用和好用, 而其理论能否健康发展在很大程度上要看是否能有效地得到计算机的支持, 即是否有优秀的计算机算法和优质的软件来支撑. 线性系统状态空间理论在应用上有强的生命力正是可借助成套的线性代数算法和软件包来保证其可计算性, 否则也只能达到表述完美, 但并不好用的窘境.

(5) 以往的代数和线性代数往往过分追求一般化, 甚至泛化, 因此缺乏相对清爽的借助几何的叙述方式. 从有用和好用的角度考虑, 矩阵是将有限维线性空间向有限维线性空间映射的线性算子. 作为线性算子, 它的特性就与由它决定的两个子空间紧密相关, 一个是列空间, 另一个是零空间. 从 70 年代开始陆续有人用这两个子空间提供的几何为主要手段来参与论述线性代数的理论收到了很好的效果. 这使我明确了写作本书的主要方法, 特别是对其基础理论的部分.

这些认识在我开始写这本书时并不很清晰, 只是模糊地认识到控制与系统科学在过去十多年里已经发生了质的巨大变化, 我们和国际同行的差距更大了, 要能迎头赶上缩小差距. 考虑到国内业界的实际水平, 必须为其提供一本反映这些变化的基础性著作, 而且在叙述方式上也应有新的特点. 可能是对这些稍显模糊信念的坚持, 在 1979 年初由汉中回迁北京前我完成了写作的框架、主要章节的构想和写作方法, 并在大家忙于搬迁的不安定环境下完成了它的雏形《应用线性代数讲义》. 回到北京以后, 我用这个讲义在国防科技大学、西北工业大学、西安交通大学、南京理工大学和成都科技大学等高校和研究所讲授并听取意见和建议, 然后利用北京更好的条件和教学与研究的实际进行充实、补充和修改, 最后在科学出版社于 1984 年春天以《系统与控制理论中的线性代数》正式出版. 清华大学自动化系对此书的问世十分重视, 在这一年的秋天我应邀到他们那里为研究生开设了这一学位课程.

结束在清华大学讲授这一课程时, 我刻意听取了一下这些青年学子的意见和建议, 他们普遍的反应是: "该课程起点高, 相当难, 一开始不适应, 后来就好了. 以前国外杂志上的文章很难看懂, 经过这半年学习, 现在有了根本性的变化." 这恰好符合我写这本书的初衷. 随后国内的一些知名院校相继以此为研究生教材或主要参考书; 一些留学生出国时将其携带身边, 以供不时之需; 甚至有些华裔教授从留

学生处见到此书后，借回国访问时也询问或来索要此书. 起初我还能满足这些朋友的需要，后来由于该书出版时还是用铅字排印，没有电子版，经过多年，且其中出版社又曾搬迁，致使铅字纸型已不存在，自 1990 年第三次印刷后即成绝版，我也只能爱莫能助了.

多年来，国内一些高校一直采用此书作为研究生教材，迫于无处可买只能自行复印或胶印以满足需求. 国内外不少学者建议该书能修改补充再版，或至少能再次重印，我皆因工作量大望而却步，并寄厚望于常用此书教学的青年才俊，但久久未能如愿. 最近李忠奎博士勇于担当组织博士生将此书重新录入形成电子版，获科学出版社鼎力支持打算再版，并希望我做适当修改以符合近 30 年发展之需要. 30 年前，改革开放迎来百废俱新的时期，当时年富力强精力充沛之人现已步入耄耋之年，再看原书，有些连自己都颇费思索感到吃力. 大改动必然力不从心，也只能在原有内容上按章节做模块式修改，即按节补充或删去整节的内容，必要时做一些补充或说明，对于这 30 年来涌现出的新的理论方法或在原书撰写时尚未认识到的有价值的内容则以独立成节的方式增补在相应章内，删补的原则以原书出版后学科发展状况为准.

过去的 30 年中，系统与控制理论有了巨大的发展，但从与线性代数有关的核心问题归纳起来主要有：

鲁棒控制的出现应该是 80 年代控制理论发生的极为重要的事件. 它主要有两个学派：一个是 Zames 提出的 H 无穷控制，在 Doyle 等的研究下将问题的解归结为两个 Riccati 矩阵方程的解，而其求解又与一类 Hamilton 矩阵的性质有关；另一个则是由俄国人 Kharitonov 的工作推动起来的参数不确定性的鲁棒分析，可归结为多项式集合的根分布问题.

如何用规划的方法研究控制一直是受人关注的问题，这方面的工作在 90 年代由于大量控制问题可以归结为线性矩阵不等式 (LMI) 的求解而得到很大发展，但是作为优化方法中将约束条件合并到指标中作统一考虑而引入的乘子方法在由不等式描述的规划问题中是否会引起增解或亏解的问题也引起了研究，另外作为一个基本优化工具的最小二乘问题也在考虑存在不确定性下取得了新进展.

早年的大系统和现在的网络的共同特点是规模大，但关注的问题和系统的结构却不同. 在对它们进行研究时，碰到两类特殊的矩阵——非负矩阵与 M 矩阵，对它们性质的讨论成了解决问题的核心，而研究这两类矩阵性质的方法又有着独到的特色，并且这两类矩阵对于研究经济系统来说确实十分重要.

对线性系统的描述常用频域与时域两种方式，两种方式之间的沟通和争论对于控制理论的发展起到了很好的促进作用. 频域语言常带有明确的物理或工程意义，而反映系统这类特性在用时域模式表述时总用矩阵、矩阵等式或不等式的性质来进行刻画，这样就引发出利用矩阵关系讨论系统性质的兴趣，而由此得到的结论常

常是可以有线性代数算法支持的.

这 30 年科技界的重大变化之一是计算机的普及, 大量基础的线性代数算法已十分丰富且自成系统, 同时发展出了相当完善的软件包. 如果说 30 年前将一些基础的算法写在书中还是必要的话, 时至今日, 在本书中继续保留相关内容就显得多余, 对于有关数值线性代数的内容, 本次修改只保留其基本原理的部分.

30 年来, 系统与控制理论的变化是巨大的, 发表的论文数量也是惊人的, 由于线性代数已在业界十分普及因此用它来研究问题的文章也一定很多, 上述只是笔者的认识, 决定修改和增补的出发点是基于以下的考虑:

① 在系统与控制理论的发展中是重要的, 并且线性代数在其中起核心作用并有新意.

② 对于最基本的常系数线性系统来说, 系统的性质与控制器的设计大都与线性代数的理论与方法有关, 其中一些也具相当难度, 而在有关控制著作中却刻意回避, 这里作了论述以利读者.

③ 考虑到当今使用计算机进行计算在业界的普及, 所有关于算法与对应的算法与程序编写的框架均删除.

30 年系统与控制理论的发展同样告诉人们: 在线性代数中有些原以为会对控制与系统学科很有用的内容可能并未显示其重要性或本身并不重要; 有些内容过于数学化与控制关系虽有但不紧; 有些内容由于学科的发展已为后来更具活力的内容所代替, 即这些只是发展进程中的不完善的中间产物等. 对于原书中这类叙述, 我将进行模块式删除.

有了这些原则, 我就可以顺当地进行模块式的修改与增补, 具体来说这一版与原书的区别在于:

第二章增加了关于多项式和多项式族根分布的一些基本结果, 这对系统的稳定性和鲁棒稳定性十分重要.

第七章增加了线性矩阵方程的基本理论.

第八章作为奇异值分解的应用, 增加了关于系统模型方面的两项内容, 模型降阶和按不同置信度分层建模.

改动最大的是第九章~第十二章, 保留了这里的基础理论, 删除了全部算法与程序的框架, 重新组合内容写成两章, 即最小二乘与优化和消元方法与特征值计算.

原书的第十三章稳定性分析现已改为第十一章.

原书的第十四章已改为第十二章, 且重新定名为有理函数矩阵与系统描述, 并增加了如下内容: H 无穷范数、全通与内稳定; 谱分解; 系统的正实性与正实引理; 小增益定理; H 无穷上的互质分解和系统镇定. 这基本上概括了常系数线性系统近 30 多年在基本描述和理论上最重要的线性代数内容.

最后增加了第十三章, 主要阐述由于系统与控制理论的发展而带动起来的一些

特殊矩阵问题. 包括: 非负矩阵; M 矩阵; 与非负矩阵相联系的一些矩阵; Hamilton 矩阵; 规划求解 (S 过程) 的无亏问题; 线性矩阵不等式及其可解性和应用: 二次稳定与 KYP 引理. 这大体上概括了近半个世纪以来在系统与控制理论发展过程中一些最具影响的方面.

考虑到运筹与控制关系密切, 近 20 年来利用规划与优化的方法解决一些控制与系统中的理论与计算的问题日益显示其优越性, 但规划与优化本身已经形成一个大的学科, 结合我们的基本需求, 写了一个 “凸性, 锥优化与对偶” 放在附录中以便于查用.

本书还对原书的保留部分作了校正, 对少数失误作了修改.

对于参考文献的选取, 我将设法按照 “必须” 和 “可追源” 两个原则, 前者是为读者理解本书提供帮助与佐证, 而后者则主要指出其出处, 这里要说明的是对于一些文献, 我只能用经过其他学者所著的有很好影响的书与综述文章来代替一批原始的文章, 这样做不仅可以让读者能由此找出源头, 也可避免大量文献的堆砌, 而且其中一些文献可能已难以查找. 对于上册, 将在每章的最后列出最基本的参考文献的索引标记, 而对下册则列在每节的后面. 书的最后是参考文献的汇总, 这里为有兴趣的读者留下较大的参考空间.

控制科学是技术科学中运用数学工具最多的学科之一, 其中的线性代数问题常不单纯地以传统线性代数理论框架中问题的形式表述. 由于系统与控制的需求而具有新的特色, 此时就不能指望仅靠基本的线性代数工具能加以解决而必须借助其他数学工具, 如分析数学、凸分析、复变函数与积分变换、线性泛函等, 这既是不可避免的, 也是符合客观发展规律的. 多种数学工具结合来解决控制与系统科学中的矩阵问题几乎已成为一种规律. 这样在本书的叙述中就必然要用其他数学而不可能也无必要保持代数的纯粹, 这可能也是学科互相渗透的必然.

本书这次出版改为分上下两册出版, 两册篇幅大致相当. 上册共八章基本上为基础理论, 不同于线性代数通常内容: 前四章基础理论涵盖了线性代数的基础理论, 但更为系统深入, 特别结合现代科学, 尤其是近数十年控制与系统理论发展的需求而成为带基础性的内容; 后四章阐述了线性代数与其他数学结合或由于其理论本身发展深化而形成的特殊理论问题, 例如矩阵范数及其应用、矩阵的广义逆及投影算子、矩阵函数, 以及在应用层面上十分奏效的奇异值分解与应用. 这两部分内容大致各占一半. 下册共五章为应用部分, 大致分三个方面: 前两章是数值代数的基础, 这是计算方法的根基之一, 而现代计算机的威力强大根本在于算法的优势. 随后的两章是现代系统与控制理论中具基础意义的部分涉及的线性代数扩展出的内容, 无论是稳定性分析, 还是有理函数矩阵的理论都是现代控制理论所必需的. 最后一章是一些特殊的矩阵类和线性矩阵不等式, 这些内容有些源于经济系统和系统工程, 有些则为用运筹与规划的理论与方法解决控制问题所必须, 而运筹与控制的结合正

是现代控制系统理论发展的必然.

年近八旬改写书稿, 当然会有困难. 一开始我并不敢为之, 但我有一个长期与我合作相对年轻的圈子, 王金枝、段志生、杨莹和李忠奎这几位教授, 他们与我相处长的 20 多年, 短的也有 10 年, 在资料汇集、组织录入, 以及最后校订等方面给我支持. 特别要提到的是李忠奎, 由于他正用此书进行教学, 就首先承担了组织博士生将此书录入成电子版的工作, 进而建议我在电子版上作适当修改, 这自然要方便得多. 由于他一方面对电子排版技术很熟悉, 同时对内容也相对清楚, 这样他就自觉地担负起在我认真校改后的在电子版上的修改工作, 从而起到了保证质量的作用. 王金枝是数学专业出身, 我在写书过程中有时会碰到一些数学上是否严谨的问题, 常会听取她的意见. 我的家庭和我的很多挚友也给我极大支持与鼓励, 在此一并表示诚挚的谢意. 相信经过努力此书的完成将无愧于业界的支持与厚望.

本书虽然经过一再努力, 但不妥之处仍在所难免, 敬请批评指正.

第一版序

现代系统与控制理论产生于 20 世纪 50 年代末期, Pontryagin 和 Bellman 关于最优控制理论方面的贡献, Kalman 关于控制系统一般理论的工作, 都是这方面的杰出代表. 从此开始的现代系统与控制理论, 一方面由于对其研究必须引进更多更深入的数学; 另一方面为了将理论应用于实际而必然与计算和计算机相结合, 这两点都使矩阵或线性代数的理论和方法在这一领域内的作用日益显著. 在五十年代的控制理论中, 一般只在少数场合才用线性代数的理论处理问题, 今天, 这种状况已经完全改观了. 现在当人们打开任何一本多变量线性系统理论的著作(即使是一本教科书)时, 都会发现在系统的描述, 系统理论中一些命题的提出、论述、解决的方法乃至结论, 都与线性代数中的概念、方法和结论紧密相联. 这种巨大的变化有其内在的深刻原因.

在系统理论模型的建立上, 线性有限维的模式总是基本的, 不仅由于这一模式便于从数学上进行处理, 而且在相当广泛的范围内这一模式可以反映系统的实质. 研究有限维线性系统最基本的工具乃是矩阵或线性代数. 即使是非线性系统的研究, 除了需要引入适应非线性特征的一些概念与方法外, 在描述与推理过程中矩阵仍然是不可缺少的手段. 就分布参数系统来说, 虽然其本质上应归于无穷维模型, 但在实际应用中常可借助一些特定的手段(差分格式或有限元), 将其近似地转化成有限维系统来进行讨论. 这种转化不仅便于使用计算机而且能较好地在实际上反映系统的性能. 这种对分布参数系统近似模型的研究手段仍然以线性代数为主.

从对 Kalman 所提出的系统可控性与可观测性的研究上, 容易发现这两个刻画系统性能的基本概念深刻地反映在关于线性变换的循环不变子空间及其生成元的论述中. 无论是系统中的各种解耦问题还是观测器理论, 也总是同矩阵方程、矩阵或线性映射所形成的各种子空间的相互关系联系在一起. 系统稳定性及二次性能指标最优的讨论, 往往归结为线性或二次矩阵代数方程或矩阵微分方程的讨论. 有时候在线性代数的领域内, 长期并未得到很好应用的结果, 在近年来控制与系统理论的推动下也获得了新的动力. 例如关于奇异值和奇异值分解的研究在今天关于系统的灵敏性分析、优化计算乃至图像识别中都有了很好的应用.

由于计算方法和计算机的发展与普及, 不仅为系统与控制理论的实际应用开辟了广阔的前景, 而且也使系统理论的研究发生着变化. 控制系统的计算机辅助设计就是适应这种变化而形成的一个分支. 由于系统与控制理论目前的发展状况, 数值线性代数中的各种理论和方法, 在相当一段时间内都会在控制与系统理论的计算机

辅助设计上起着重要作用.

用线性代数的基本理论来处理系统与控制理论中的问题, 往往描述简洁而且便于抓住实质. 这一点已为近二十年来的事实所证实. 例如当采用线性变换的几何理论来讨论状态空间型系统时, 就易于把握住问题的核心而得到理论上深刻的结论, 甚至像商空间这样抽象的概念也能在系统与控制理论上发挥重要作用. 在用算子的多项式矩阵或有理函数矩阵描述的系统的研究中, 多项式矩阵环及其上的理想的理论起着重要的作用. 这方面的工作不仅已经通过 Laurent 多项式矩阵环的理论推广至更为一般的线性系统, 而且它还可以与经典的控制理论相联系起到独特的作用. 为了寻求对线性系统更普遍的描述方式与研究工具, 一些人从代数中较抽象的一个分支模论出发进行研究, 也已经显示出一种新的图景.

从上面的分析可以断言: 线性代数的理论与方法是研究现代系统与控制理论的重要数学基础. 本书正是基于这一要求而撰写的.

本书共四个部分:

第一部分是线性代数的基本理论. 它包括线性空间与线性映射; 多项式与多项式矩阵; 线性变换的代数理论与几何理论, 酉空间与正规矩阵等. 叙述采用空间的概念与矩阵形式相结合的方法, 以求简洁与概括. 阅读这部分内容不仅可以达到整理与加深理解常规线性代数理论的目的, 而且也为今后的展开创造一个良好的条件. 此外, 在结合系统与控制理论的需要上也作了必要的展开, 例如关于多项式矩阵的理想、线性变换的结构特别关于循环不变子空间及其生成元的论述都是为此目的而写出的.

第二部分是线性代数的几个特殊理论问题, 它包括矩阵范数和它的应用; 矩阵的摄动理论; 矩阵函数; 广义逆与投影算子理论; 矩阵的奇异值分解以及极小化理论等. 这部分内容的选取不仅考虑到当前系统与控制理论发展的状况, 而且也为今后发展的需要做了一定准备.

第三部分是数值线性代数中与系统和控制理论关系密切的部分, 它包括线性方程组的直接法; 无约束与有约束的线性最小二乘解以及矩阵的特征值计算. 这一部分的叙述在给出严谨的理论论证的同时还给出了一种粗线条的原则性的算法, 以期为应用计算机进行计算创造一定条件. 对于基于差分法解偏微分方程的需要而发展起来的各种矩阵迭代方法, 限于篇幅将不作任何论述.

第四部分讲了两个专门问题: 其中一个是与稳定性及二次型最优相联系的矩阵代数方程问题, 由于引进了矩阵的 Kronecker 乘积而带来了方便, 对于 Lyapunov 方程、代数 Riccati 方程以及 Hurwitz 问题均作了讨论. 另一个是关于系统矩阵及有理函数矩阵的内容, 这一部分对于多变量系统的讨论是很必要的.

本书每章均留有一些练习, 一些比较简单的命题 (包括引理、定理与推论) 未给出证明, 这些均提供读者以练习的机会. 附录给出了书中符号的统一规定和代数

的必备的基本知识, 以便读者查阅. 书后列有必要的参考文献, 由于文献浩瀚, 不可能周全, 一些已见诸于书的成果就不再求源引原文献而只引有关书籍, 以保证读者能有线索可寻.

本书的前身是我写的一本《线性代数应用理论讲义》, 该讲义的部分内容先后在近十个高等院校及研究所讲过, 本书是在这些讲学基础上联系到系统和控制理论的需要改写而成的. 在撰写本书的过程中, 我得到了很多单位及同志们的支持和帮助, 其中特别应该提到的有关肇直、宋健、高为炳、张志方、秦化淑、贺建勋、于景元、郑应平和王恩平等同志, 他们或给予作者以热情支持或在内容与写法等方面提供了宝贵意见.

北京大学力学系特别是一般力学教研室的同志给予作者以热情的鼓励和支持, 北京大学计算数学教研室王颖坚同志阅读了有关数值线性代数的章节并提供了不少有益的意见.

对于在撰写本书过程中曾给予作者以支持和帮助的所有同志和单位, 在此一并表示深切的感谢.

限于水平, 书中不当乃至错误之处难免, 热忱欢迎批评指正, 以期今后改进.

目　录

第九章　最小二乘问题

最小化问题是规划理论也是系统理论中的一个基本问题, 这里讨论的是可以用线性代数手段来进行讨论的最小化问题的理论部分. 这里主要讨论无约束最小二乘解问题, 具一次或凸二次约束的最小二乘解问题, 以及当系数矩阵出现不确定性时的最小二乘问题.

由于最小化问题的解决总与广义逆矩阵的讨论有关, 因而这里将充分使用广义逆, 投影算子作为基本工具.

9.1　最小二乘解问题及其基本理论结果

在线性代数中一个基本问题是求解方程组

$$Ax = b, \tag{9.1.1}$$

其中 $A \in \mathbf{C}^{m\times n}$, $b \in \mathbf{C}^m$. 对于式 (9.1.1) 其可解性条件归结为是否有

$$b \in \mathbf{R}(A). \tag{9.1.2}$$

在相当一类实际问题中, 条件 (9.1.2) 常常不能满足而要求研究:

最小二乘解问题 (LS 问题): 给定 $A \in \mathbf{C}^{m\times n}$, $b \in \mathbf{C}^m$, 要求求解 x 使

$$\|Ax - b\|_2 = \min. \tag{9.1.3}$$

一般称上述问题为无约束最小二乘解问题, 如果在 $\mathbf{C}^n$ 中给定一个描述约束的集合 $\mathbf{S}$, 则称问题

$$\|Ax - b\|_2 = \min, \quad x \in \mathbf{S} \tag{9.1.4}$$

为具约束 $\mathbf{S}$ 的 LS 问题.

定理 9.1.1　$A \in \mathbf{C}^{m\times n}$, $b \in \mathbf{C}^m$ 则问题 (9.1.3) 当 $x = A^{[1,3]}b$ 时达到极小, 其中 $A^{[1,3]} \in \mathbf{A}\{1,3\}$. 反之, 若 $X \in \mathbf{C}^{n\times m}$ 能对一切 b 有 $\|Ax-b\|_2$ 在 $x = Xb$ 时达到极小, 则 $X \in \mathbf{A}\{1,3\}$.

证明　令 $b = b_1 + b_2$, 其中 $b_1 = P_{\mathbf{R}(A)}b$, $b_2 = (I - P_{\mathbf{R}(A)})b$, 则有

$$\|Ax - b\|_2^2 = \|Ax - b_1\|_2^2 + \|b_2\|_2^2 \geqslant \|b_2\|_2^2. \tag{9.1.5}$$

利用式 (9.1.5) 可知 x 是式 (9.1.3) 的解当且仅当

$$Ax = b_1 = P_{\mathbf{R}(A)}b = AA^{[1,3]}b. \tag{9.1.6}$$

由此可知任何 $X \in \mathbf{A}\{1,3\}$, 都有 $x = Xb$ 是式 (9.1.6) 从而是式 (9.1.3) 的解.

反之若 X 对任何 $b \in \mathbf{C}^m$ 均使 $x = Xb$ 实现 $\|Ax - b\|_2 = \min$, 则特别依次选 $b = e_1, e_2, \cdots, e_n$, 对应 Xe_i 应满足

$$AXe_i = AA^{[1,3]}e_i, \quad i \in \underline{n},$$

而这相当于 X 满足

$$AX = AA^{[1,3]},$$

显然必有 $X \in \mathbf{A}\{1,3\}$. ■

容易证明有下述推论.

推论 9.1.1　LS 问题 (9.1.3) 的通解是

$$x = A^{[1,3]}b + \mathbf{N}(A). \tag{9.1.7}$$

因而, LS 问题 (9.1.3) 的解唯一当且仅当

$$\mathbf{N}(A) = \{0\}. \tag{9.1.8}$$

■

定理 9.1.2　x 是式 (9.1.3) 的解当且仅当满足

$$A^HAx = A^Hb, \tag{9.1.9}$$

并且式 (9.1.3) 的解对应的余量 $r = Ax - b \in \mathbf{N}(A^H)$.

证明　x 是式 (9.1.3) 的解 $\Longleftrightarrow Ax = P_{\mathbf{R}(A)}b \Longleftrightarrow Ax - b = -(I - P_{\mathbf{R}(A)})b \in \mathbf{N}(A^H) \Longleftrightarrow A^H(Ax - b) = 0 \Longleftrightarrow x$ 有式 (9.1.9).

x 是式 (9.1.3) 的解则余量 $Ax - b \in \mathbf{N}(A^H)$ 为显然. ■

一般称式 (9.1.9) 为问题 (9.1.3) 的正规方程.

显然, 对任何 $A \in \mathbf{C}^{m\times n}$, $b \in \mathbf{C}^m$ 来说正规方程 (9.1.9) 总有解.

推论 9.1.2　若 $A = FG$ 分解中 G 具满行秩, 则方程 (9.1.9) 等价于方程

$$F^HAx = F^Hb. \tag{9.1.10}$$

■

推论 9.1.3　x 是式 (9.1.9) 的解当且仅当有 r 使

$$\begin{bmatrix} -I & A \\ A^H & 0 \end{bmatrix}\begin{bmatrix} r \\ x \end{bmatrix} = \begin{bmatrix} b \\ 0 \end{bmatrix}. \tag{9.1.11}$$

■