现代数学基础

21 控制论中的矩阵计算

■ 徐树方 著

高等教育出版社
HIGHER EDUCATION PRESS

21

控制论中的矩阵计算

■ 徐树方　著

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

内容简介

本书主要介绍控制论中几个典型矩阵计算问题的数值解法。全书共分7章,内容包括:矩阵分析基础、控制系统概论、矩阵指数的计算、Lyapunov方程的数值解法、代数Riccati方程的数值解法、非对称代数Riccati方程的数值解法、极点配置问题的数值解法。本书在内容上,力求向读者展示这一领域既基本又重要的知识、方法和技巧以及最新的进展。本书在叙述表达上,力求清晰易读,便于教学与自学。

本书可作为综合大学、理工科大学、高等师范院校计算数学、应用数学、工程计算等专业高年级本科生和研究生的教材或教学参考书,也可供从事科学与工程计算的科技人员参考。

图书在版编目(CIP)数据

控制论中的矩阵计算/徐树方著.—北京:高等教育出版社,2011.3
ISBN 978-7-04-031611-7

Ⅰ.①控… Ⅱ.①徐… Ⅲ.①矩阵-计算方法-高等学校-教材 Ⅳ.①O241.6

中国版本图书馆CIP数据核字(2011)第002117号

策划编辑	赵天夫	**责任编辑**	蒋　青	**封面设计**	赵　阳
版式设计	余　杨	**责任校对**	刘　莉	**责任印制**	刘思涵

出版发行	高等教育出版社	**购书热线**	010-58581118
社　　址	北京市西城区德外大街4号	**咨询电话**	400-810-0598
邮政编码	100120	**网　　址**	http://www.hep.edu.cn
			http://www.hep.com.cn
		网上订购	http://www.landraco.com
经　　销	蓝色畅想图书发行有限公司		http://www.landraco.com.cn
印　　刷	北京中科印刷有限公司	**畅想教育**	http://www.widedu.com
开　　本	787×1092　1/16	**版　　次**	2011年3月第1版
印　　张	23	**印　　次**	2011年3月第1次印刷
字　　数	430 000	**定　　价**	59.00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。

物料号　31611-00

前言

在 20 世纪 90 年代初, 我将自己在科研中所收集的资料进行了整理和加工, 在北京大学数学科学学院科学与工程计算系给高年级本科生和研究生开设了一个以控制论中的若干矩阵计算问题为主要讲授内容的选修课。在这之后的近二十年的教学和科研中, 我又不断地对有关的内容进行了更新和增补, 使所讲授的内容日渐丰富和成熟。我觉得所讲授的这些方法和理论对从事科学与工程计算的科技人员是有益的, 因此近两年来集中精力将其整理成书, 以飨读者。

本书共分七章。第一章对本书常用的一些矩阵分析中的基本概念和重要结果作了概述性的介绍, 第二章给出了本书所涉及的控制论中的一些基本概念和重要结果, 后面的五章对控制论中的五个最基本的矩阵计算问题进行了深入细致的讨论, 内容包括: 矩阵指数的计算、Lyapunov 方程的数值解法、代数 Riccati 方程的数值解法、非对称代数 Riccati 方程的数值解法、极点配置问题的数值解法。本书在材料的安排上, 是按照计算数学科研活动的基本规律安排的, 即对每个问题, 先建立其可解性理论, 再分析其解对扰动的敏感性, 最后给出求解它的数值方法。本书在内容的处理上, 力求向读者展示这一领域最基本的知识、最重要的方法和技巧以及最新的进展, 尽可能地阐明每一数值方法的设计思想和理论依据, 并且尽可能地对每个结论给出严格而又简洁的数学证明。本书在叙述表达上, 力求清晰易读, 便于教学与自学。此外, 书中每章都安排了几道难易程度不同的习题, 并且对该章主要结果和方法的来源以及进一步探究的一些线索作了简要的说明。书末还附有较详细的参考文献目录。

北京大学数学科学学院科学与工程计算系的研究生刘站营和卢欣曾帮助我

将部分书稿录入计算机, 我的太太连惠民女士帮助我做了全书的图并且录入了部分书稿, 中国矿业大学的刘兰冬老师曾仔细阅读了全部书稿并指出了原稿中存在的不少错误, 高等教育出版社的策划编辑赵天夫和责任编辑蒋青为本书的出版付出了辛勤的劳动, 在此一并表示诚挚的谢意。

欢迎读者对本书提出批评指正。

徐树方

2010 年 8 月

目录

前言

第一章　矩阵分析基础

为了阅读本书方便起见, 这一章我们将本书常用的一些矩阵分析中的基本概念和重要结果作一概述性的介绍. 这里假定读者已经熟悉有限维向量空间的基本概念和初等性质.

§1.1　基本概念和常用符号

为了以后行文方便和避免不必要的重复, 这一节将本书中常用的一些符号和术语作一简要的说明.

本书中, 我们约定用小写希腊字母 α, β, ξ, η 等表示实数或复数, 实数的全体用 $\mathscr{R}$ 表示, 而复数的全体用 $\mathscr{C}$ 表示; 用小写英文字母 i, j, k, l, m, n 等表示整数; 用小写英文字母 a, b, x, y 等表示实向量或复向量, 实 n 维列向量的全体用 $\mathscr{R}^n$ 表示, 而复 n 维列向量的全体用 $\mathscr{C}^n$ 来表示; 用大写英文字母 A, B, X, Y 等来表示矩阵, 实 $m \times n$ 矩阵的全体用 $\mathscr{R}^{m\times n}$ 来表示, 而复 $m \times n$ 矩阵的全体用 $\mathscr{C}^{m\times n}$ 来表示.

对于给定的 $A \in \mathscr{C}^{m\times n}$, 我们用 $A^\top, \bar{A}$ 和 A^* 分别表示矩阵 A 的转置矩阵、共轭矩阵和共轭转置矩阵, 而用 $\operatorname{rank} A$ 表示矩阵 A 的秩. 如果 $m = n$, 则简称 A 为 n 阶矩阵, 并且用 $\det A$ 和 $\operatorname{tr} A$ 分别表示 A 的行列式和迹.

如果矩阵 $A \in \mathscr{C}^{n\times n}$ 满足 $\det A \neq 0$, 则称 A 为**非奇异的**, 否则称 A 为**奇异的**. 如果 A 非奇异, 则用 A^{-1} 来表示 A 的逆矩阵.

n 阶单位矩阵用 I_n 来表示, I_n 的第 k 列记作 $e_k^{(n)}$, 即

$$I_n = \left[e_1^{(n)}, e_2^{(n)}, \ldots, e_n^{(n)}\right],$$
$$e_k^{(n)} = (0, \ldots, 0, \underset{k}{1}, 0, \ldots, 0)^{\top} \in \mathscr{C}^n.$$

当维数从上下文中一目了然时, 简记为 I 和 e_k. 对单位矩阵 I 的列进行重新排列之后所得到的矩阵称为 n 阶**排列方阵**, 即排列方阵是每列、每行只有唯一的一个非零元素 "1" 的矩阵.

设 $A = [\alpha_{ij}] \in \mathscr{C}^{m\times n}$, 如果 A 满足

$$\alpha_{ij} = 0, \qquad i < j,$$

则称 A 是**下三角形矩阵**, 对角元均为 1 的下三角形矩阵称作**单位下三角形矩阵**. 如果 A 满足

$$\alpha_{ij} = 0, \qquad i > j,$$

则称 A 是**上三角形矩阵**, 对角元均为零的下三角形矩阵, 称作**严格下三角形矩阵**; 对角元均为零的上三角形矩阵, 称作**严格上三角形矩阵**. 若 A 满足

$$\alpha_{ij} = 0, \qquad i \neq j,$$

则称 A 是**对角矩阵**. 我们常用 $\mathrm{diag}\,(\alpha_1, \alpha_2, \ldots, \alpha_n)$ 来表示对角元为 $\alpha_1, \alpha_2, \ldots, \alpha_n$ 的 n 阶对角矩阵. 当 α_j 被方阵代替后, 它表示块对角矩阵.

设 $A \in \mathscr{C}^{n\times n}$, 如果 $AA^* = A^*A$, 则称 A 是**正规矩阵**; 如果 $A^* = A$, 则称 A 是 **Hermite 矩阵**; 如果 $A^* = -A$, 则称 A 是**反 Hermite 矩阵**. 如果 $A^*A = AA^* = I$, 则称 A 是**酉矩阵**. 显然, Hermite 矩阵、反 Hermite 矩阵和酉矩阵都是正规矩阵. 这里值得注意的是并非所有的正规矩阵必是这三类之一. 例如 $A = (i+1)I$ 是正规矩阵, 但它并不属于这三类矩阵的任一类. 实的 Hermite 矩阵、实的反 Hermite 矩阵和实的酉矩阵分别称为**对称矩阵**、**反对称矩阵**和**正交矩阵**. $n \times n$ 实对称矩阵的全体记作 $\mathscr{R}_s^{n\times n}$. 对于一个 Hermite 矩阵 A, 如果对任意的非零向量 $x \in \mathscr{C}^n$ 都有 $x^*Ax > 0$, 则称 A 是**正定的**, 记作 $A \succ 0$; 如果对任意的向量 $x \in \mathscr{C}^n$ 都有 $x^*Ax \geqslant 0$, 则称 A 是**半正定的**, 记作 $A \succcurlyeq 0$.

设 $\mathscr{X}$ 是 $\mathscr{C}^n$ 的一个子空间, 我们用 $\dim \mathscr{X}$ 来表示 $\mathscr{X}$ 的维数, 而用 $\mathscr{X}^{\perp}$ 来表示 $\mathscr{X}$ 的正交补, 即

$$\mathscr{X}^{\perp} = \{y \in \mathscr{C}^n \ : \ y \perp \mathscr{X}\},$$

这里 $y \perp \mathscr{X}$ 是指 y 与 $\mathscr{X}$ 中的任一向量都正交, 即 $y^*x = 0$ 对一切的 $x \in \mathscr{X}$ 都成立. 特别地, 对任意给定的非零向量 $x \in \mathscr{C}^n$, 定义

$$x^{\perp} = \{y \in \mathscr{C}^n \ : \ y^*x = 0\},$$

称之为向量 x 的**正交补空间**.

设 $x_1,x_2,\ldots,x_k$ 是 $\mathscr{C}^n$ 中的给定的 k 个向量, 我们用 $\mathrm{span}\{x_1,x_2,\ldots,x_k\}$ 来表示由此 k 个向量张成的子空间, 即

$$\mathrm{span}\{x_1,x_2,\ldots,x_k\}=\left\{\sum_{j=1}^{k}\alpha_j x_j\ :\ \alpha_j\in\mathscr{C}\right\}.$$

$\dim\mathscr{X}=k$ 意指 $\mathscr{X}$ 中存在 k 个线性无关的向量 $x_1,x_2,\ldots,x_k$, 使得 $\mathscr{X}=\mathrm{span}\{x_1,x_2,\ldots,x_k\}$.

对于任一给定的 $A\in\mathscr{C}^{m\times n}$, 有两个与 A 有关的重要的子空间: 一个是 A 的**列空间**, 定义为

$$\mathscr{R}(A)=\{Ax\ :\ x\in\mathscr{C}^n\}\subset\mathscr{C}^m,$$

另一个是 A 的**零空间**, 定义为

$$\mathscr{N}(A)=\{x\in\mathscr{C}^n\ :\ Ax=0\}\subset\mathscr{C}^n.$$

易知, $\mathscr{R}(A)$ 就是由 A 的列向量张成的子空间.

当我们限定在实数范围内讨论问题时, 上述定义也要限定在实数范围内. 例如, 若 $x_1,x_2,\ldots,x_k$ 是 $\mathscr{R}^n$ 中的向量, 则

$$\mathrm{span}\{x_1,x_2,\ldots,x_k\}=\left\{\sum_{j=1}^{k}\alpha_j x_j\ :\ \alpha_j\in\mathscr{R}\right\}.$$

再如, 若 $A\in\mathscr{R}^{m\times n}$, 则有

$$\mathscr{N}(A)=\{x\in\mathscr{R}^n\ :\ Ax=0\}\subset\mathscr{R}^n.$$

令 $\iota=(i_1,i_2,\ldots,i_k)$, 其中 i_j 是正整数, 且 $1\leqslant i_1<i_2<\cdots<i_k\leqslant n$. 然后记全体这样的 ι 所构成的集合为 $\mathscr{Q}_{kn}$. 设 $A=[\alpha_{ij}]\in\mathscr{C}^{n\times n}$, 对任意的 $\iota\in\mathscr{Q}_{kn}$, 记

$$A(\iota)=\begin{bmatrix}\alpha_{i_1i_1} & \alpha_{i_1i_2} & \cdots & \alpha_{i_1i_k}\\ \alpha_{i_2i_1} & \alpha_{i_2i_2} & \cdots & \alpha_{i_2i_k}\\ \vdots & \vdots & & \vdots\\ \alpha_{i_ki_1} & \alpha_{i_ki_2} & \cdots & \alpha_{i_ki_k}\end{bmatrix},$$

并称之为 A 的一个 k **阶主子阵**. 特别, 当 $\iota=(1,2,\ldots,k)$ 时, 对应的 k 阶主子阵 $A(\iota)$ 被称为 A 的 k **阶顺序主子阵**.

设

$$A(t)=\begin{bmatrix}\alpha_{11}(t) & \alpha_{12}(t) & \cdots & \alpha_{1n}(t)\\ \alpha_{21}(t) & \alpha_{22}(t) & \cdots & \alpha_{2n}(t)\\ \vdots & \vdots & & \vdots\\ \alpha_{m1}(t) & \alpha_{m2}(t) & \cdots & \alpha_{mn}(t)\end{bmatrix}\in\mathscr{R}^{m\times n},$$

其中 $\alpha_{ij}(t)$ 是实变元 t 的函数. 我们用 $\dot{A}(t)$ 和 $\int_{t_0}^{t}A(\tau)d\tau$ 分别表示 $A(t)$ 的元逐个进行微分和积分所得到的矩阵, 即

$$\dot{A}(t)=\begin{bmatrix}\dfrac{d\alpha_{11}(t)}{dt} & \dfrac{d\alpha_{12}(t)}{dt} & \cdots & \dfrac{d\alpha_{1n}(t)}{dt}\\ \dfrac{d\alpha_{21}(t)}{dt} & \dfrac{d\alpha_{22}(t)}{dt} & \cdots & \dfrac{d\alpha_{2n}(t)}{dt}\\ \vdots & \vdots & & \vdots\\ \dfrac{d\alpha_{m1}(t)}{dt} & \dfrac{d\alpha_{m2}(t)}{dt} & \cdots & \dfrac{d\alpha_{mn}(t)}{dt}\end{bmatrix},$$

$$\int_{t_0}^{t}A(\tau)d\tau=\begin{bmatrix}\int_{t_0}^{t}\alpha_{11}(\tau)d\tau & \int_{t_0}^{t}\alpha_{12}(\tau)d\tau & \cdots & \int_{t_0}^{t}\alpha_{1n}(\tau)d\tau\\ \int_{t_0}^{t}\alpha_{21}(\tau)d\tau & \int_{t_0}^{t}\alpha_{22}(\tau)d\tau & \cdots & \int_{t_0}^{t}\alpha_{2n}(\tau)d\tau\\ \vdots & \vdots & & \vdots\\ \int_{t_0}^{t}\alpha_{m1}(\tau)d\tau & \int_{t_0}^{t}\alpha_{m2}(\tau)d\tau & \cdots & \int_{t_0}^{t}\alpha_{mn}(\tau)d\tau\end{bmatrix}.$$

特别地, 当 $x(t)=\big(x_1(t),x_2(t),\dots,x_n(t)\big)^{\top}\in\mathscr{R}^n$ 时, 有

$$\dot{x}(t)=\big(x_1'(t),x_2'(t),\dots,x_n'(t)\big)^{\top},$$

$$\int_{t_0}^{t}x(\tau)d\tau=\left(\int_{t_0}^{t}x_1(\tau)d\tau,\int_{t_0}^{t}x_2(\tau)d\tau,\dots,\int_{t_0}^{t}x_n(\tau)d\tau\right)^{\top}.$$

此外, 我们也常用 $k\ll n$ 来表示整数 k 远远小于整数 n.

§1.2 初等矩阵及其应用

大家知道, 在线性代数中初等变换起到非常重要的作用. 这一节我们就对这些初等变换作一综述性的介绍, 并给出其中几个最基本的应用.

§1.2.1 初等矩阵

一、初等矩阵的一般形式

所谓的**初等矩阵**是指如下形式的矩阵

$$E(u,v;\sigma)=I-\sigma uv^*, \tag{1.2.1}$$

其中 $u,v\in\mathscr{C}^n$, $\sigma\in\mathscr{C}$.

根据初等矩阵的定义, 容易验证它有如下性质:

(1) $\det E(u,v;\sigma)=1-\sigma v^*u$;

(2) $E(u,v;\sigma)$ 非奇异的充要条件是 $\sigma v^*u\neq 1$, 且当其非奇异时, 有

$$E(u,v;\sigma)^{-1}=E\big(u,v;\sigma(\sigma v^*u-1)^{-1}\big); \tag{1.2.2}$$

(3) 对 $\mathscr{C}^n$ 中的任意两个非零向量 a 和 b, 必可适当选取 $u,v\in\mathscr{C}^n$ 和 $\sigma\in\mathscr{C}$, 使得

$$E(u,v;\sigma)a=b. \tag{1.2.3}$$

注 1.2.1 事实上, 对于 (1.2.3), 我们只需取 u,v 和 σ 满足

$$v^*a\neq 0,\quad \sigma u=\frac{a-b}{v^*a} \tag{1.2.4}$$

即可.

顺便指出, 线性代数中的三种初等变换, 都可用现在定义的初等矩阵给出:

- 交换一个矩阵的第 i 行和第 j 行等价于对该矩阵左乘一个初等矩阵

$$E(e_i-e_j,e_i-e_j;1)=I-(e_i-e_j)(e_i-e_j)^\top.$$

- 将一个矩阵的第 i 行乘以 α 再加到第 j 行等价于对该矩阵左乘一个初等矩阵

$$E(e_j,e_i;-\alpha)=I+\alpha e_je_i^\top.$$

- 一个矩阵的第 i 行乘以 α 等价于对该矩阵左乘一个初等矩阵

$$E(e_i,e_i;1-\alpha)=I-(1-\alpha)e_ie_i^\top.$$

下面我们着重讨论初等下三角形矩阵和初等 Hermite 矩阵, 它们是矩阵计算的最基本工具.

二、初等下三角形矩阵

初等下三角形矩阵是指形如

$$L_j(l_j) \equiv E(l_j, e_j; 1) = I - l_j e_j^\top \tag{1.2.5}$$

的初等矩阵, 其中

$$l_j = (0, \ldots, 0, \zeta_{j+1,j}, \ldots, \zeta_{nj})^\top . \tag{1.2.6}$$

初等下三角形矩阵又称为 **Gauss 变换**. 根据初等下三角形矩阵的定义和初等矩阵的性质, 易知初等下三角形矩阵有如下性质:

(1) $\det L_j(l_j) = 1$;

(2) $L_j(l_j)^{-1} = L_j(-l_j)$;

(3) 对任意给定的 $a = (\alpha_1, \ldots, \alpha_n)^\top$ $(\alpha_j \neq 0)$, 可确定唯一的初等下三角形矩阵 $L_j(l_j)$, 使得

$$L_j(l_j)a = (\alpha_1, \ldots, \alpha_j, 0, \ldots, 0)^\top, \tag{1.2.7}$$

其中 l_j 的非零分量由下式给出

$$\zeta_{ij} = \frac{\alpha_i}{\alpha_j}, \quad i = j+1, \ldots, n. \tag{1.2.8}$$

此外, 若 $i < j$, 则有 $(l_i e_i^\top)(l_j e_j^\top) = 0$. 利用这一性质, 立即可知, 任一单位下三角形矩阵

$$L = \begin{bmatrix} 1 & & & & \\ -\zeta_{21} & 1 & & & \\ -\zeta_{31} & -\zeta_{32} & \ddots & & \\ \vdots & \vdots & \ddots & \ddots & \\ -\zeta_{n1} & -\zeta_{n2} & \cdots & -\zeta_{n,n-1} & 1 \end{bmatrix}$$

必可分解为

$$L = L_1(l_1)L_2(l_2)\cdots L_{n-1}(l_{n-1}),$$

其中 l_j 如 (1.2.6) 所示.

三、初等 Hermite 矩阵

初等 Hermite 矩阵是在 (1.2.1) 中取 $v = u = w$, $w^*w = 1$, $\sigma = 2$ 所得到的初等矩阵, 即初等 Hermite 矩阵有如下形式

$$H(w) \equiv E(w, w; 2) = I - 2ww^*, \tag{1.2.9}$$

其中 $w \in \mathscr{C}^n$ 满足 $w^*w = 1$. 在矩阵计算中, 常将上述初等 Hermite 矩阵称作 **Householder 变换**.

根据初等 Hermite 矩阵的定义和初等矩阵的性质, 容易验证初等 Hermite 矩阵具有如下基本性质:

(1) $\det H(w) = -1$;

(2) $H(w)^* = H(w) = H(w)^{-1}$;

(3) 对任一 $u \in w^{\perp}$, 有

$$H(w)(u + \beta w) = u - \beta w, \qquad \beta \in \mathscr{C};$$

(4) 设 $a, b \in \mathscr{C}^n$ $(a \neq b)$, 则存在向量 $w \in \mathscr{C}^n$ 满足 $w^*w = 1$, 使得 $H(w)a = b$ 的充分与必要条件是

$$a^*a = b^*b \quad 和 \quad a^*b = b^*a, \tag{1.2.10}$$

并且在 (1.2.10) 所示的条件成立时, 所需的向量 w 可取作

$$w = e^{i\theta} \frac{a - b}{\sqrt{(a-b)^*(a-b)}}, \tag{1.2.11}$$

其中 θ 为任一实数.

注 1.2.2 初等 Hermite 矩阵的性质 (3) 的几何解释如图 1.2.1 所示, 即 $H(w)(u + \beta w)$ 是 $(u + \beta w)$ 关于 w 的垂直超平面的镜像反射.

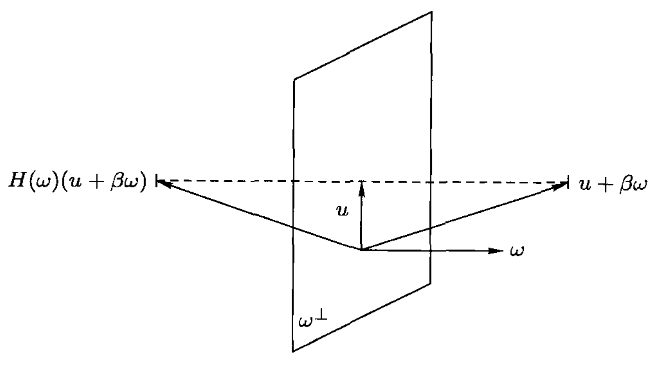

图 1.2.1 镜像反射

注 1.2.3 由初等 Hermite 矩阵的性质 (4) 可知, 对任给的

$$a = (\alpha_1, \alpha_2, \ldots, \alpha_n)^{\top} \neq 0,$$