Introduction to Adaptive Control of Non-linear Systems

非线性系统的自适应控制导论

周东华 编著

Zhou Donghua

TUP
清华大学出版社

Springer
施普林格出版社

(京)新登字 158 号

内 容 简 介

本书主要介绍线性系统自适应控制的典型方法，以及基于强跟踪滤波器理论的非线性系统的自适应控制，基于模糊集理论的非线性系统的自适应控制，基于神经元网络理论的非线性系统的自适应控制，鲁棒自适应控制的基本知识，以及非线性系统的鲁棒自适应控制方法。

本书适于作为机电控制类的研究生教材，同时对从事自动化系统的研究、设计、开发和应用的广大工程技术人员也有一定的参考价值。

书 名：非线性系统的自适应控制导论
作 者：周东华 编著
出版者：清华大学出版社 施普林格出版社
北京清华大学学研大厦，邮编 100084
http：//www. tup. tsinghua. edu. cn
印刷者：清华大学印刷厂
发行者：新华书店总店北京发行所
开 本：787×1092 1/16 印张：10.5 字数：217 千字
版 次：2002 年 1 月第 1 版 2002 年 1 月第 1 次印刷
书 号：ISBN 7-302-04479-1/TP・2643
印 数：0001～3000
定 价：18.00 元

前言

自适应控制的发展已有 40 多年的历史，并且在近 20 多年里得到了飞速的发展，已成为当代自动控制界的少数热门前沿研究领域之一。大批的科技工作者在从事自适应控制的研究与应用开发工作，每年都有数千篇的学术论文与研究报告问世。自适应控制理论已经在航空航天、机器人、冶金、造纸、啤酒酿造、航海、水电站、机车控制、化工、窑炉控制、水下勘探等众多的工程技术领域得到了成功的应用，取得了显著的社会效益与经济效益。

任何自适应控制系统都是非线性系统。自适应控制的主要理论基础是随机过程、线性代数、系统辨识、稳定性理论等。一些新兴的理论与技术，如：模糊集理论、人工神经元网络、小波变换和遗传算法等，对自适应控制理论的发展也起到了很大的推动作用。

我们在清华大学自动化系研究生课程的教学过程中发现，国内还没有一本比较系统性地介绍非线性系统自适应控制的教材，而非线性系统的自适应控制已成为当前自适应控制领域的热点研究方向。因此在机电类研究生中开设非线性系统的自适应控制课程已具有一定的迫切性，这正是我们撰写这本教材的源动力之一。我们深信，本书的出版将会对我国非线性系统的自适应控制理论的教学与科研工作起到一定的促进作用。

作为一本非线性系统自适应控制的导论性教材，本书共分为 7 章。第 1 章介绍一些基础知识，包括：特殊矩阵；L_p 空间，范数及矩阵的几种重要分解；动态系统的稳定性；指数稳定性定理；正实函数；超稳定性的基本概念等。第 2 章介绍一些线性系统自适应控制的典型方法，包括：最小方差自校正调节器，广义最小方差控制器，PID 自整定控制器，基于 Lyapunov 稳定性理论的模型参考自适应控制器，以及基于超稳定性理论的模型参考自适应控制器的基本内容。这些内容构成了经典线性系统自适应控制的最核心部分。第 3 章介绍基于强跟踪滤波器理论的非线性系统的自适应控制方法，这也是作者近年来的研究成果。第 4 章介绍基于模糊集理论的非线性系统的自适应控制方法。第 5 章给出

基于神经元网络理论的非线性系统的自适应控制方法。第6章介绍鲁棒自适应控制的基础知识,包括:鲁棒性的基本概念,Rohrs的关于存在未建模动态时的自适应算法的鲁棒性问题,以及吴宏鑫教授等提出的黄金分割法在鲁棒自适应控制器设计中的应用。本书的最后一章介绍三种非线性系统的鲁棒自适应控制方法,分别是:具有结构非线性扰动的一类动态系统的鲁棒自适应采样数据控制,具有测量干扰的基于神经网络的非线性系统的鲁棒自适应控制,以及机械臂的鲁棒自适应控制。

本书的主要内容都选自近年来发表在国际著名学术刊物上的代表性文章。因此,本书也适用于立志从事非线性系统自适应控制理论研究的学者作为导论性的读物。由于时间仓促和编者的学术水平有限,书中肯定会有不少错误,恳请广大读者批评指正。

本书第3章所介绍的研究成果曾得到了国家自然科学基金会和教育部有关项目的资助,本书的出版还得到了清华大学985学科建设经费的资助,在此表示衷心的感谢。

周东华

2001年3月

于清华园

目录

第1章 基 础 知 识

本章介绍本书用到的一些基础知识，包括：特殊矩阵；L_p 空间，范数及矩阵的几种重要分解；动态系统的稳定性；指数稳定性定理；正实函数；超稳定性的基本概念等。

1.1 记号

标量和向量用小写字母表示。

矩阵、算子、集合用大写字母表示。

$u(s)$表示 $u(t)$的拉氏变换。

线性时不变系统，简记为 LTI(Linear Time Invariant)。

LTI 系统的传递函数用 $H(s)$表示。

1.2 特殊矩阵

定义 1.2.1 一个方阵 $\boldsymbol{A}\in R^{n\times n}$是半正定的(positive semi-definite)，如果 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}\geqslant 0$ 对所有的向量 $\boldsymbol{x}$ 成立.

定义 1.2.2 一个方阵 $\boldsymbol{A}\in R^{n\times n}$是正定的(positive definite)，如果 $\exists\alpha>0$，使得 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}\geqslant\alpha\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=\alpha|\boldsymbol{x}|^2$ 对所有的 $\boldsymbol{x}$ 成立。

半正定矩阵的特征值都位于右半闭平面(closed right half plane, RHP)。正定矩阵的特征值都位于右半开平面。

定义 1.2.3 若 $\boldsymbol{A}\geqslant 0$,并且 $\boldsymbol{A}=\boldsymbol{A}^{\mathrm{T}}$,那么称 $\boldsymbol{A}$ 是对称的半正定阵。对称矩阵的特征值都是实数,这样的矩阵有 n 个正交的特征向量。因此我们可以分解 $\boldsymbol{A}$ 为

$$\boldsymbol{A}=\boldsymbol{U}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{U} \tag{1.2.1}$$

其中,$\boldsymbol{U}$ 是特征向量矩阵,满足 $\boldsymbol{U}^{\mathrm{T}}\boldsymbol{U}=\boldsymbol{I}$。$\boldsymbol{\Lambda}$ 是对角矩阵,对角元素正是 $\boldsymbol{A}$ 的 n 个特征值。

定义 1.2.4 平方根矩阵 $\boldsymbol{\Lambda}^{1/2}$ 是对角矩阵,其对角元素是由对称矩阵 $\boldsymbol{A}$ 的特征根的平方根构成,并且有

$$\boldsymbol{A}^{1/2}=\boldsymbol{U}^{\mathrm{T}}\boldsymbol{\Lambda}^{1/2}\boldsymbol{U} \tag{1.2.2}$$

是 $\boldsymbol{A}$ 的平方根阵,即

$$\boldsymbol{A}=\boldsymbol{A}^{1/2}\boldsymbol{A}^{1/2},\qquad (\boldsymbol{A}^{1/2})^{\mathrm{T}}=\boldsymbol{A}^{1/2}$$

若 $\boldsymbol{A}\geqslant 0$, $\boldsymbol{B}\geqslant 0$,则必有 $\boldsymbol{A}+\boldsymbol{B}\geqslant 0$,但是,$\boldsymbol{AB}\geqslant 0$ 并不一定成立。

若 $\boldsymbol{A},\boldsymbol{B}$ 是对称的半正定阵,那么 $\boldsymbol{AB}$ 虽然不一定是对称阵,但 $\boldsymbol{AB}$ 的特征值都是实数。

对称、半正定阵 $\boldsymbol{A}$ 的另一个特性是

$$\lambda_{\min}(\boldsymbol{A})\,|\boldsymbol{x}|^2\leqslant \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}\leqslant \lambda_{\max}(\boldsymbol{A})\,|\boldsymbol{x}|^2 \tag{1.2.3}$$

$$\|\boldsymbol{A}\|=\lambda_{\max}(\boldsymbol{A}) \tag{1.2.4}$$

若 $\boldsymbol{A}$ 是正定的,则又有

$$\|\boldsymbol{A}^{-1}\|=1/\lambda_{\min}(\boldsymbol{A}) \tag{1.2.5}$$

定义 1.2.5 称矩阵 $\boldsymbol{A}$ 为幂等矩阵,若 $\boldsymbol{A}^2=\boldsymbol{A}$。

命题 1.2.1 任何一个幂等矩阵都是半正定的[3]。

定义 1.2.6 满足条件 $\boldsymbol{A}^{\mathrm{T}}=-\boldsymbol{A}$ 的正方矩阵称为反对称矩阵或斜对称矩阵。显然,为了满足反对称性,主对角线上的元素必须为零,并且有

$$a_{ij}=-a_{ji},\quad i\neq j \tag{1.2.6}$$

反对称矩阵具有以下性质:

(1) 若 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都是反对称矩阵,则 $\boldsymbol{A}^{-1}$,$\boldsymbol{A}+\boldsymbol{B}$ 仍是反对称矩阵。并且当 k 为偶数时,$\boldsymbol{A}^k$ 为对称矩阵;当 k 为奇数时,$\boldsymbol{A}^k$ 为反对称矩阵[3]。

(2) 任意正方矩阵 $\boldsymbol{A}$ 都可以分解为一个对称矩阵 $\frac{1}{2}(\boldsymbol{A}+\boldsymbol{A}^{*})$ 和一个反对称矩阵 $\frac{1}{2}(\boldsymbol{A}-\boldsymbol{A}^{*})$ 之和。

定义 1.2.7 交换矩阵 $\boldsymbol{J}$ 定义如下:

$$\boldsymbol{J}=\begin{bmatrix}0 & \cdots & 0 & 1\\ 0 & \cdots & 1 & 0\\ \vdots & \ddots & \vdots & \vdots\\ 1 & \cdots & 0 & 0\end{bmatrix} \tag{1.2.7}$$

它仅在交叉对角线上具有元素 1，而所有其他元素全等于零。$\boldsymbol{J}$ 矩阵可以使一个矩阵的行或列的顺序反转（互换）。如用 $\boldsymbol{J}_m$ 左乘 $m\times n$ 的矩阵 $\boldsymbol{A}$，将使 $\boldsymbol{A}$ 的行的顺序反转：

$$\boldsymbol{J}_m\boldsymbol{A}=\begin{bmatrix} a_{m1} & a_{m2} & \cdots & a_{mn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{21} & a_{22} & \cdots & a_{2n} \\ a_{11} & a_{12} & \cdots & a_{1n} \end{bmatrix} \tag{1.2.8}$$

容易验证，$\boldsymbol{J}^{\mathrm{T}}=\boldsymbol{J}$，$\boldsymbol{J}^2=\boldsymbol{J}\boldsymbol{J}=\boldsymbol{I}$。

定义 1.2.8 我们称向量 $\boldsymbol{x}_1,\cdots,\boldsymbol{x}_k\in C^m$ 组成一正交组，若 $\boldsymbol{x}_i^{\mathrm{H}}\boldsymbol{x}_j=0\ (1\leqslant i<j\leqslant k)$；另外，若这些向量还是归一化的，即 $\boldsymbol{x}_i^{\mathrm{H}}\boldsymbol{x}_i=1$，$i=1,2,\cdots,k$，则称该正交组是标准正交的。

定义 1.2.9 一个实的方阵 $\boldsymbol{A}\in R^{n\times n}$ 称为正交矩阵，若 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{I}$；一个复值方阵 $\boldsymbol{U}\in C^{n\times n}$ 称为酉矩阵，若 $\boldsymbol{U}^{\mathrm{H}}\boldsymbol{U}=\boldsymbol{I}$。其中，$\boldsymbol{U}^{\mathrm{H}}$ 表示矩阵 $\boldsymbol{U}$ 的共轭转置。

命题 1.2.2（酉矩阵的性质） 对矩阵 $\boldsymbol{U}\in R^{n\times n}$，下面的叙述是等价的[3]：

（1）$\boldsymbol{U}$ 是酉矩阵；

（2）$\boldsymbol{U}$ 是非奇异的，并且 $\boldsymbol{U}^{\mathrm{H}}=\boldsymbol{U}^{-1}$；

（3）$\boldsymbol{U}\boldsymbol{U}^{\mathrm{H}}=\boldsymbol{I}$；

（4）$\boldsymbol{U}^{\mathrm{H}}$ 是酉矩阵；

（5）$\boldsymbol{U}$ 的列组成标准正交组；

（6）$\boldsymbol{U}$ 的行组成标准正交组；

（7）对所有的 $\boldsymbol{x}\in C^n$，$\boldsymbol{y}=\boldsymbol{U}\boldsymbol{x}$ 的 Euclidean 长度与 $\boldsymbol{x}$ 的 Euclidean 长度相等，即 $\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}=\boldsymbol{y}^{\mathrm{H}}\boldsymbol{y}$。

定义 1.2.10 若 $\boldsymbol{U}$ 是酉矩阵，一个满足 $\boldsymbol{B}=\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}$ 的矩阵 $\boldsymbol{B}\in C^{n\times n}$ 被称为是与 $\boldsymbol{A}\in R^{n\times n}$ 酉等价的，也可以说 $\boldsymbol{A}$ 酉相似于 $\boldsymbol{B}$。

定义 1.2.11 对矩阵 $\boldsymbol{A}=[a_{ij}]\in C^{n\times n}$，若 $\boldsymbol{A}=\boldsymbol{A}^{\mathrm{H}}$，则称其为 Hermitian 矩阵，其中，$\boldsymbol{A}^{\mathrm{H}}=(\boldsymbol{A}^*)^{\mathrm{T}}=[a_{ij}^*]$；若 $\boldsymbol{A}=-\boldsymbol{A}^{\mathrm{H}}$，则称矩阵 $\boldsymbol{A}$ 为斜 Hermitian 的或反 Hermitian 矩阵。

Hermitian 矩阵具有如下性质：

（1）对所有 $\boldsymbol{A}\in C^{n\times n}$，矩阵 $\boldsymbol{A}+\boldsymbol{A}^{\mathrm{H}}$，$\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}$ 和 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ 均是 Hermitian 的。

（2）若 $\boldsymbol{A}$ 是 Hermitian 阵，则 $\boldsymbol{A}^k$ 对所有 $k=1,2,3,\cdots$ 都是 Hermitian 阵；若 $\boldsymbol{A}$ 还是非奇异的，则 $\boldsymbol{A}^{-1}$ 也是 Hermitian 阵。

（3）若矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是 Hermitian 的，则 $\alpha\boldsymbol{A}+\beta\boldsymbol{B}$ 对所有的实数 α 和 β 均是 Hermitian 的。

（4）$\boldsymbol{A}-\boldsymbol{A}^{\mathrm{H}}$ 对所有的 $\boldsymbol{A}\in C^{n\times n}$ 是反 Hermitian 的。

（5）若矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是反 Hermitian 的，则 $\alpha\boldsymbol{A}+\beta\boldsymbol{B}$ 对所有的实数 α 和 β 均是反 Hermitian 的。

（6）若 $\boldsymbol{A}$ 是 Hermitian 的，则 $\mathrm{j}\boldsymbol{A}$（其中，$\mathrm{j}=\sqrt{-1}$）是反 Hermitian 的。

（7）若 $\boldsymbol{A}$ 是反 Hermitian 的，则 $\mathrm{j}\boldsymbol{A}$ 是 Hermitian 的。

(8) 为一个方阵,并且其对角元为实。

(9) 其特征值为实数。

(10) 若 $\boldsymbol{A}$ 为 Hermitian 矩阵,$\boldsymbol{x}$ 为一个具有复数分量的矢量,则二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}^{*}$ 恒为实。

定理 1.2.1(Hermitian 矩阵的谱定理)[3] 令 $\boldsymbol{A}\in C^{n\times n}$ 已知,则 $\boldsymbol{A}$ 是 Hermitian 矩阵的充要条件是,存在一个酉矩阵 $\boldsymbol{U}\in C^{n\times n}$ 和一个实的对角矩阵 $\boldsymbol{\Lambda}\in R^{n\times n}$,使得:

$$\boldsymbol{A}=\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^{\mathrm{H}} \tag{1.2.9}$$

$\boldsymbol{A}$ 是实的 Hermitian(实对称)矩阵的充要条件是,存在实的正交矩阵 $\boldsymbol{P}\in R^{n\times n}$ 和一个实的对角矩阵 $\boldsymbol{\Lambda}\in R^{n\times n}$,使得

$$\boldsymbol{A}=\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{\mathrm{T}} \tag{1.2.10}$$

假定 $\boldsymbol{A}$ 的 n 个特征向量为 q_i, $i=1,2,\cdots,n$,它的 n 个特征值为 λ_i, $i=1,2,\cdots,n$。那么(1.2.9)式的另一种表示形式是

$$\boldsymbol{A}=\sum_{i=1}^{n}\lambda_i\boldsymbol{q}_i\boldsymbol{q}_i^{\mathrm{H}} \tag{1.2.11}$$

式(1.2.11)也称作矩阵 $\boldsymbol{A}$ 的谱分解。

由关系式

$$\boldsymbol{A}^{-1}=(\boldsymbol{U}^{\mathrm{H}})^{-1}\boldsymbol{\Lambda}^{-1}\boldsymbol{U}^{-1}=\boldsymbol{U}\boldsymbol{\Lambda}^{-1}\boldsymbol{U}^{\mathrm{H}}$$

我们有

$$\boldsymbol{A}^{-1}=\sum_{i=1}^{n}\frac{1}{\lambda_i}\boldsymbol{q}_i\boldsymbol{q}_i^{\mathrm{H}} \tag{1.2.12}$$

这里假定所有的特征值都不等于零。

定义 1.2.12 满足条件 $a_{ij}=0$, $i>j$ 的正方矩阵 $\boldsymbol{A}=[a_{ij}]$ 称为上三角矩阵,其一般形式为

$$\boldsymbol{A}=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22} & \cdots & a_{2n} \\ & & \ddots & \vdots \\ & & & a_{nn} \end{bmatrix} \tag{1.2.13}$$

特别地,若 $a_{ii}>0$, $i=1,2,\cdots,n$, 则称矩阵 $\boldsymbol{A}$ 为正线上三角矩阵。

1.3 L_p 空间、范数及矩阵的几种重要分解

如果 x 是标量,则 $|x|$ 表示 x 的绝对值;若 $\boldsymbol{x}$ 是向量,$|\boldsymbol{x}|$ 则表示 $\boldsymbol{x}$ 的欧氏范数(l_2 范数)。

$\|\cdot\|$ 一般用来表示算子的诱导范数,如矩阵的诱导范数:

$$\|A\| = \sup_{|x|=1} |Ax| \tag{1.3.1}$$

对时间函数，此记号表示 L_p 范数：

$$\|u\|_p = \left(\int_0^{\infty} |u(\tau)|^p \mathrm{d}\tau\right)^{1/p} \tag{1.3.2}$$

对 $p\in[1,\infty)$成立。其中，定义：

$$\|u\|_{\infty} = \sup_{t\geqslant 0} |u(t)| \tag{1.3.3}$$

定义 1.3.1 若 $\|u\|_p$ 存在，我们称 $u\in L_p$ 空间。当 p 省略时，$\|u\|$ 表示 L_2 范数，即

$$\|u\| = \|u\|_2 = \sqrt{\int_0^{\infty} u^2(\tau)\mathrm{d}\tau}$$

定义 1.3.2 截头函数定义为

$$f_s(t) = \begin{cases} f(t), & t\leqslant s \\ 0, & t > s \end{cases} \tag{1.3.4}$$

定义 1.3.3 扩展的 L_p 空间定义为

$$L_{pe} = \{f \mid \forall s < \infty,\ f_s \in L_p\} \tag{1.3.5}$$

如：e^t 不属于L_{∞}，但 $e^t\in L_{\infty e}$空间。

当 $u\in L_{\infty e}$时，有

$$\|u_t\|_{\infty} = \sup_{\tau\leqslant t} |u(\tau)| \tag{1.3.6}$$

可以证明[1]：若函数 $f\in L_1\cap L_{\infty}$，则 $\forall p\in[1,\infty)$，$f\in L_p$。

引理 1.3.1(Barbalat's Lemma) 若 $f(t)$是一致连续函数，并且 $\lim\limits_{t\to\infty}\int_0^t f(\tau)\mathrm{d}\tau$ 存在且有极限值，那么

$$\lim_{t\to\infty} f(t) \to 0$$

推论 1.3.1 如果 g，$\dot{g}\in L_{\infty}$，并且 $\exists p\in[1,\infty)$使得 $g\in L_p$，那么 $\lim\limits_{t\to\infty} g(t)\to 0$。

1.3.1 矩阵的 UR 分解及其推论

我们首先探讨满秩方阵的 UR 分解问题。

引理 1.3.2[4] 每一个 n 维内积空间 V 一定存在标准正交基底。

定理 1.3.1 设 $A\in C_n^{n\times n}$，则存在酉矩阵 $U\in C^{n\times n}$ 及正线上三角矩阵 $R\in C_n^{n\times n}$，使得 $A=UR$。其中，$C_n^{n\times n}$ 的右下标 n 表示矩阵 A 的秩。

证明：因为 $A\in C_n^{n\times n}$，故 A 为满秩方阵，将 A 写成列分块形式：

$$A = [a_1, a_2, \cdots, a_n]$$

则 a_i，$i=1,2,\cdots,n$ 均为 C^n 向量，且线性无关，因此 a_i，$i=1,2,\cdots,n$ 可作为 C^n 的一组正交基底。根据引理 1.3.2，用这组基底可以求出 C^n 的一组标准正交基底：

$$\{\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n\}$$

并可以表示成[4]：

$$[\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n] = [\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n]\boldsymbol{R}$$

其中，$\boldsymbol{R}$ 为正线上三角型。

把上式写成矩阵形式，即有

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{R} \tag{1.3.7}$$

注意到 $\boldsymbol{U}=[\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n]$的各列标准正交，故 $\boldsymbol{U}^{\mathrm{H}}\boldsymbol{U}=\boldsymbol{I}$，即 $\boldsymbol{U}$ 为一个酉矩阵。

[证毕]

推论 1.3.2 设 $\boldsymbol{A}\in R_n^{n\times n}$，则存在正交矩阵 $\boldsymbol{Q}\in R^{n\times n}$ 及正线上三角矩阵 $\boldsymbol{R}\in R_n^{n\times n}$，使得 $\boldsymbol{A}=\boldsymbol{Q}\boldsymbol{R}$。

下面，我们进一步介绍长方矩阵的分解问题。

定理 1.3.2 设 $\boldsymbol{A}\in C_n^{m\times n}$，$m>n$，则存在酉矩阵 $\boldsymbol{U}\in C^{m\times m}$ 及正线上三角矩阵 $\boldsymbol{R}_1$，使得

$$\boldsymbol{A} = \boldsymbol{U}\begin{bmatrix}\boldsymbol{R}_1\\ \boldsymbol{O}\end{bmatrix} \tag{1.3.8}$$

其中，$\boldsymbol{O}\in C^{(m-n)\times n}$，$\boldsymbol{R}_1\in C_n^{n\times n}$。

证明[4]：记 $\boldsymbol{A}=[\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n]$，则 $\boldsymbol{a}_i\in C^m$，$i=1,2,\cdots,n$。又因为 $\boldsymbol{A}$ 为列满秩，故 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n$ 为 C^m 中线性无关向量组，将它扩充为 C^m 的基底，记为

$$\{\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n, \boldsymbol{a}_{n+1}, \cdots, \boldsymbol{a}_m\}$$

由引理 1.3.2 得知，可将这组基底标准化，得到 C^m 中的一组标准正交基底：

$$\{\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_m\}$$

且有正线上三角形 $\boldsymbol{R}$，使得

$$[\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n, \boldsymbol{a}_{n+1}, \cdots, \boldsymbol{a}_m] = [\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_m]\boldsymbol{R} \tag{1.3.9}$$

将 $\boldsymbol{R}$ 分块表示，有

$$\boldsymbol{R} = \begin{bmatrix}\boldsymbol{R}_1 & \boldsymbol{C}\\ \boldsymbol{O} & \boldsymbol{R}_2\end{bmatrix}\in C^{m\times n} \tag{1.3.10}$$

其中，$\boldsymbol{R}_1\in C^{n\times n}$，$\boldsymbol{C}\in C^{n\times(m-n)}$，$\boldsymbol{O}\in C^{(m-n)\times n}$，$\boldsymbol{R}_2\in C^{(m-n)\times(m-n)}$，且 $\boldsymbol{R}_1$，$\boldsymbol{R}_2$ 均为正线上三角形矩阵。将(1.3.10)式代入(1.3.9)式，并由分块矩阵相乘的性质得

$$[\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n] = [\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_m]\begin{bmatrix}\boldsymbol{R}_1\\ \boldsymbol{O}\end{bmatrix} \tag{1.3.11}$$

即

$$\boldsymbol{A} = \boldsymbol{U}\begin{bmatrix}\boldsymbol{R}_1\\ \boldsymbol{O}\end{bmatrix} \tag{1.3.12}$$

其中，$\boldsymbol{U}$ 显然为酉矩阵。

[证毕]

推论 1.3.3[4] 设 $\boldsymbol{A}\in C_r^{m\times n}$，则存在酉矩阵 $\boldsymbol{U}\in C^{m\times m}$ 和酉矩阵 $\boldsymbol{V}\in C^{n\times n}$，以及正线上三角阵 $\boldsymbol{R}\in C_r^{r\times r}$，使得

$$\boldsymbol{A}=\boldsymbol{U}\begin{bmatrix}\boldsymbol{R} & \boldsymbol{O}_{r\times(n-r)}\\ \boldsymbol{O}_{(m-r)\times r} & \boldsymbol{O}_{(n-r)\times(m-r)}\end{bmatrix}\boldsymbol{V}^{\mathrm{H}} \tag{1.3.13}$$

推论 1.3.4 设 $\boldsymbol{A}$ 为一个实对称正定矩阵，则存在正线上三角矩阵 $\boldsymbol{R}$，使得 $\boldsymbol{A}=\boldsymbol{R}^{\mathrm{T}}\boldsymbol{R}$。

证明：因为 $\boldsymbol{A}$ 为对称正定矩阵，因此，$\boldsymbol{A}$ 可以经过一系列行变换和一系列同名的列变换化为单位矩阵，即存在满秩矩阵 $\boldsymbol{P}$，使得 $\boldsymbol{A}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}$。又由推论 1.3.2 对 $\boldsymbol{P}$ 存在正交矩阵 $\boldsymbol{Q}$ 及正线上三角矩阵 $\boldsymbol{R}$，使得 $\boldsymbol{P}=\boldsymbol{Q}\boldsymbol{R}$，因此

$$\boldsymbol{A}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}=\boldsymbol{R}^{\mathrm{T}}\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{R}=\boldsymbol{R}^{\mathrm{T}}\boldsymbol{R} \tag{1.3.14}$$

[证毕]

定义 1.3.4 称定理 1.3.1 给出的分解式为满秩方阵的 UR 分解；称推论 1.3.3 给出的分解为长方阵的正交分解；称推论 1.3.4 给出的分解为 Cholesky 分解或三角-三角分解。

定理 1.3.3(满秩分解定理)[4] 设 $\boldsymbol{A}\in C_r^{m\times n}$，则存在列满秩矩阵 $\boldsymbol{B}\in C_r^{m\times r}$ 和行满秩矩阵 $\boldsymbol{C}\in C_r^{r\times n}$，使得

$$\boldsymbol{A}=\boldsymbol{B}\boldsymbol{C} \tag{1.3.15}$$

1.3.2 舒尔引理与正规矩阵的分解

定理 1.3.4(舒尔引理)[4] 设 $\boldsymbol{A}\in C^{n\times n}$，则存在酉矩阵 $\boldsymbol{U}\in C^{n\times n}$ 及上三角矩阵 $\boldsymbol{T}=[t_{ij}]\in C^{n\times n}$，使得

$$\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}=\boldsymbol{T} \tag{1.3.16}$$

且 $\boldsymbol{T}$ 的主对角元素均为 $\boldsymbol{A}$ 的特征值。

定义 1.3.5 若 $\boldsymbol{A}\in C^{n\times n}$满足条件：

$$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}} \tag{1.3.17}$$

则称 $\boldsymbol{A}$ 为正规矩阵。

由定义可知，下列方阵均为正规矩阵：

(1) 实对称矩阵($\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}$, $\boldsymbol{A}\in R^{n\times n}$)；

(2) 实反对称矩阵($\boldsymbol{A}^{\mathrm{T}}=-\boldsymbol{A}$, $\boldsymbol{A}\in R^{n\times n}$)；

(3) Hermitian 矩阵($\boldsymbol{A}^{\mathrm{H}}=\boldsymbol{A}$, $\boldsymbol{A}\in C^{n\times n}$)；

(4) 反 Hermitian 矩阵($\boldsymbol{A}^{\mathrm{H}}=-\boldsymbol{A}$, $\boldsymbol{A}\in C^{n\times n}$)；

(5) 正交矩阵($\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}^{-1}$, $\boldsymbol{A}\in R^{n\times n}$)；

(6) 酉矩阵($\boldsymbol{U}^{\mathrm{H}}=\boldsymbol{U}^{-1}$, $\boldsymbol{U}\in C^{n\times n}$)。

正规矩阵具有如下重要性质：

定理 1.3.5 设 $\boldsymbol{A}\in C^{n\times n}$，则 $\boldsymbol{A}$ 是正规矩阵的充要条件为：存在一个 n 阶酉矩阵 $\boldsymbol{U}$ 及

一个 n 阶对角矩阵 $\boldsymbol{\Lambda}=\mathrm{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$，使得

$$\boldsymbol{A}=\boldsymbol{U\Lambda U}^{\mathrm{H}} \tag{1.3.18}$$

这时称 $\boldsymbol{U\Lambda U}^{\mathrm{H}}$ 为 $\boldsymbol{A}$ 的酉相似对角化分解式，简称为 $\boldsymbol{A}$ 的酉相似对角分解。

证明：(1) 必要性

若 $\boldsymbol{A}$ 为正规矩阵，则有：

$$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}} \tag{1.3.19}$$

由舒尔引理，存在酉矩阵 $\boldsymbol{U}\in C^{n\times n}$ 及上三角矩阵 $\boldsymbol{T}\in C^{n\times n}$，使得 $\boldsymbol{A}=\boldsymbol{UTU}^{\mathrm{H}}$。将此式代入(1.3.19)式可得

$$(\boldsymbol{UTU}^{\mathrm{H}})^{\mathrm{H}}(\boldsymbol{UTU}^{\mathrm{H}})=(\boldsymbol{UTU}^{\mathrm{H}})(\boldsymbol{UTU}^{\mathrm{H}})^{\mathrm{H}}$$

注意到 $\boldsymbol{U}^{\mathrm{H}}\boldsymbol{U}=\boldsymbol{I}$，则由上式可以推出

$$\boldsymbol{T}^{\mathrm{H}}\boldsymbol{T}=\mathbf{T}\mathbf{T}^{\mathrm{H}} \tag{1.3.20}$$

令

$$\boldsymbol{T}=\begin{bmatrix} t_{11} & t_{12} & \cdots & \cdots & t_{1n} \\ 0 & t_{22} & \cdots & \cdots & t_{2n} \\ \vdots & & & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & t_{nn} \end{bmatrix} \tag{1.3.21}$$

将(1.3.21)式代入(1.3.20)式，比较等式两边矩阵对应位置上的元素，可得

$$\sum_{i=1}^{n} |t_{1i}|^2=|t_{11}|^2 \tag{1.3.22}$$

$$\sum_{i=1}^{n} |t_{2i}|^2=|t_{12}|^2+|t_{22}|^2 \tag{1.3.23}$$

$$\vdots$$

$$\sum_{i=1}^{n} |t_{ni}|^2=|t_{nn}|^2 \tag{1.3.24}$$

由(1.3.22)式得知，必有 $t_{12}=t_{13}=\cdots=t_{1n}=0$，因此，(1.3.23)式变成了 $\sum_{i=2}^{n}|t_{2i}|^2=|t_{22}|^2$，由此又可以推出 $t_{23}=t_{24}=\cdots=t_{2n}=0$。如此递推下去，可推出 $\boldsymbol{T}$ 必须为对角矩阵。令 $\boldsymbol{T}$ 的主对角线上元素 $t_{ii}(i=1,2,\cdots,n)$ 为 λ_i，并记 $\boldsymbol{\Lambda}=\mathrm{diag}\{\lambda_1,\lambda_2,\cdots,\lambda_n\}$，因此有

$$\boldsymbol{A}=\boldsymbol{UTU}^{\mathrm{H}}=\boldsymbol{U\Lambda U}^{\mathrm{H}} \tag{1.3.25}$$

(2) 充分性

若 $\boldsymbol{A}$ 酉相似于对角矩阵 $\boldsymbol{\Lambda}$，则(1.3.25)式成立。由于 $\boldsymbol{\Lambda}$ 为对角阵，因此有

$$\boldsymbol{\Lambda}^{\mathrm{H}}\boldsymbol{\Lambda}=\boldsymbol{\Lambda\Lambda}^{\mathrm{H}} \tag{1.3.26}$$

所以有

$$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=(\boldsymbol{U\Lambda U}^{\mathrm{H}})^{\mathrm{H}}(\boldsymbol{U\Lambda U}^{\mathrm{H}})=\boldsymbol{U\Lambda}^{\mathrm{H}}\boldsymbol{U}^{\mathrm{H}}\boldsymbol{U\Lambda U}^{\mathrm{H}}=\boldsymbol{U\Lambda}^{\mathrm{H}}\boldsymbol{\Lambda U}^{\mathrm{H}}=\boldsymbol{U\Lambda\Lambda}^{\mathrm{H}}\boldsymbol{U}^{\mathrm{H}}$$

$$\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}=\boldsymbol{U\Lambda U}^{\mathrm{H}}(\boldsymbol{U\Lambda U}^{\mathrm{H}})^{\mathrm{H}}=\boldsymbol{U\Lambda U}^{\mathrm{H}}\boldsymbol{U\Lambda}^{\mathrm{H}}\boldsymbol{U}^{\mathrm{H}}=\boldsymbol{U\Lambda\Lambda}^{\mathrm{H}}\boldsymbol{U}^{\mathrm{H}}$$

即 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ 成立，$\boldsymbol{A}$ 为正规矩阵。

［证毕］

定理 1.3.6（Schur 不等式）[4] 若 n 阶复方阵 $\boldsymbol{A}=(a_{ij})$ 的特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$，则有

$$\sum_{i=1}^{n}|\lambda_i|^2\leqslant\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2 \tag{1.3.27}$$

其中，等号当且仅当 $\boldsymbol{A}$ 是正规矩阵时成立。

证明：由舒尔引理，存在酉矩阵 $\boldsymbol{U}$ 及上三角阵 $\boldsymbol{T}$，使得

$$\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{U}=\boldsymbol{T} \tag{1.3.28}$$

因此有

$$\boldsymbol{T}\boldsymbol{T}^{\mathrm{H}}=\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{U} \tag{1.3.29}$$

由矩阵求迹的性质，对任意方阵 $\boldsymbol{B}$ 和 $\boldsymbol{C}$，有 $\mathrm{tr}\{\boldsymbol{BC}\}=\mathrm{tr}\{\boldsymbol{CB}\}$，因此，

$$\begin{aligned}\mathrm{tr}\{\boldsymbol{T}\boldsymbol{T}^{\mathrm{H}}\}&=\mathrm{tr}\{\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{U}\}=\mathrm{tr}\{\boldsymbol{A}^{\mathrm{H}}\boldsymbol{U}\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}\}=\mathrm{tr}\{\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\}\\&=\mathrm{tr}\{\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}\}=\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2\end{aligned} \tag{1.3.30}$$

由于 $\boldsymbol{T}$ 的主对角元就是 $\boldsymbol{A}$ 的特征值，因此有

$$\sum_{i=1}^{n}|\lambda_i|^2=\sum_{i=1}^{n}|t_{ii}|^2\leqslant\sum_{i=1}^{n}|t_{ii}|^2+\sum_{j=2}^{n}\sum_{i=1}^{j-1}|t_{ij}|^2=\mathrm{tr}\{\boldsymbol{T}\boldsymbol{T}^{\mathrm{H}}\}=\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2$$

上述不等式中的等号当且仅当 $t_{ij}=0\ (i\neq j)$ 时成立，即当且仅当 $\boldsymbol{A}$ 为正规矩阵时成立。

［证毕］

推论 1.3.5[4] 若 $\boldsymbol{A}\in R^{n\times n}$ 为实对称阵，则 $\boldsymbol{A}$ 的特征值 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 皆为实数，且存在正交阵 $\boldsymbol{Q}\in R^{n\times n}$，使得

$$\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q}=\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}=\mathrm{diag}\{\lambda_1,\lambda_2,\cdots,\lambda_n\} \tag{1.3.31}$$

1.3.3 矩阵的奇异值分解

由推论(1.3.3)得知，对每个复矩阵 $\boldsymbol{A}\in C_r^{m\times n}$ 都有分解式：

$$\boldsymbol{A}=\boldsymbol{U}\begin{bmatrix}\boldsymbol{R}&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{O}\end{bmatrix}\boldsymbol{V}^{\mathrm{H}}$$

其中，$\boldsymbol{U}$，$\boldsymbol{V}$ 分别为 m 阶和 n 阶的酉矩阵，而 $\boldsymbol{R}$ 是一个 r 阶的正线上三角矩阵。下面我们对上述形式进一步简化，将 $\boldsymbol{R}$ 化成一个对角矩阵，且主对角元取 $\boldsymbol{A}$ 的奇异值，由此得到 $\boldsymbol{A}$ 的奇异值分解。

定义 1.3.6 设 $\boldsymbol{A}\in C^{m\times n}$，则矩阵 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ 的 n 个特征值 λ_i，$i=1,2,\cdots,n$ 的算术平方根 $\sigma_i=\sqrt{\lambda_i}$ 称作 $\boldsymbol{A}$ 的奇异值。

定理 1.3.7（矩阵的奇异值分解定理） 设 $\boldsymbol{A}\in C_r^{m\times n}$，则存在酉矩阵 $\boldsymbol{U}\in C^{m\times m}$ 及 $\boldsymbol{V}\in$