矩阵函数的求法和矩阵分解方法研究

摘 要

本论文基于矩阵论课程，详细介绍了待定系数法、数项级数求和法、对角型法、若当标准型法四种矩阵函数求法，及矩阵的 LU 分解、QR 分解、满秩分解、奇异值分解四种矩阵分解方法，解释了这些方法的理论基础，并对它们分别进行了举例说明，在论文开篇补充了欧氏空间、线性变换、范数等相关预备知识。合适的矩阵函数和矩阵分解方法能够大大减少工程运算量及提高计算速度、精度。

关键词：矩阵函数　矩阵分解

Research on the solution of matrix function and matrix decomposition method

ABSTRACT

Based on the course of matrix theory, this paper introduces in detail four matrix function solving methods: undetermined coefficient method, sum of several series method, diagonal type method and Jordan Standard type method, as well as four matrix decomposition methods: LU decomposition, QR decomposition, full rank decomposition and singular value decomposition, explains the theoretical basis of these methods, and gives examples, At the beginning of the paper, the relevant preparatory knowledge such as Euclidean space, linear transformation and norm are supplemented. Appropriate matrix function and matrix decomposition method can greatly reduce the amount of engineering calculation and improve the calculation speed and accuracy.

KEY WORDS: Matrix function Matrix decomposition

目　录

第一章 引言

1.1 背景介绍

1.1.1 矩阵理论与方法介绍

在数学中，矩阵是一个按照长方阵列排列的复数或实数集合，最早来自于方程组的系数及常数所构成的方阵。这一概念由 19 世纪英国数学家凯利首先提出。作为解决线性方程的工具，矩阵也有不短的历史。成书最早在东汉前期的《九章算术》中，用分离系数法表示线性方程组，得到了其增广矩阵。在消元过程中，使用的把某行乘以某一非零实数、从某行中减去另一行等运算技巧，相当于矩阵的初等变换。但那时并没有现今理解的矩阵概念，虽然它与现有的矩阵形式上相同，但在当时只是作为线性方程组的标准表示与处理方式。

矩阵是高等代数学中的常见工具，也常见于统计分析等应用数学学科中。 在物理学中，矩阵于电路学、力学、光学和量子物理中都有应用；计算机科学中，三维动画制作也需要用到矩阵。矩阵的运算是数值分析领域的重要问题。将矩阵分解为简单矩阵的组合可以在理论和实际应用上简化矩阵的运算。对一些应用广泛而形式特殊的矩阵，例如稀疏矩阵和准对角矩阵，有特定的快速运算算法。在天体物理、量子力学等领域，也会出现无穷维的矩阵，是矩阵的一种推广。人工智能对矩阵的应用也是颇为广泛。

矩阵正式作为数学中的研究对象出现，则是在行列式的研究发展起来后。逻辑上，矩阵的概念先于行列式，但在实际的历史上则恰好相反。日本数学家关孝和与微积分的发现者之一戈特弗里德·威廉·莱布尼茨（1693 年）近乎同时地独立建立了行列式论。其后行列式作为解线性方程组的工具逐步发展。1750 年，加布里尔·克拉默发现了克莱姆法则。

矩阵的概念在 19 世纪逐渐形成。1800 年代，高斯和威廉·若尔当建立了高斯—若尔当消去法。1844 年，德国数学家费迪南·艾森斯坦讨论了“变换”（矩阵）及其乘积。1850 年，英国数学家詹姆斯·约瑟夫·西尔维斯特首先使用矩阵一词 。

英国数学家阿瑟·凯利被公认为矩阵论的奠基人。他开始将矩阵作为独立的数学对象研究时，许多与矩阵有关的性质已经在行列式的研究中被发现了，这也使得凯利认为矩阵的引进是十分自然的。他说：“我决然不是通过四元数而获得矩阵概念的；它或是直接从行列式的概念而来，或是作为一个表达线性方程组的方便方法而来的。”他从 1858 年开始，发表了《矩阵论的研究报告》等一系列关于矩阵的专门论文，研究了矩阵的运算律、矩阵的逆以及转置和特征多项式方程。凯利还提出了凯莱-哈密尔顿定理，并验证了 3×3 矩阵的情况，又说进一步的证明是不必要的。哈密尔顿证明了 4×4 矩阵的情况，而一般情况下的证明是德国数学家弗罗贝尼乌斯于 1898 年给出的。

1854 年时法国数学家埃尔米特使用了“正交矩阵”这一术语，但他的正式定义直到 1878 年才由费罗贝尼乌斯发表。1879 年，费罗贝尼乌斯引入矩阵秩的概念。至此，矩阵的体系基本上建立起来了。

矩阵理论是学习是学习经典数学的基础，又是一门最有实用价值的数学理论。它不仅是数学的一个重要的分支，而且业已成为现代各科技领域处理大量有限维空间形式与

数量关系的强有力的工 具，特别是计算机的广泛应用。为矩阵论的应用开辟了广阔的前景。例如系统工程优化方法以及稳定性理论等都与矩阵论有着密切的联系。

矩阵理论与方法介绍了在实际工程中有应用价值的矩阵理论与方法，包括线性空间与线性变换、范数理论及其应用、矩阵分析及其应用、矩阵分解、特征值的估计及对称矩阵的极性、广义逆矩阵等。

学习矩阵理论与方法对计算机专业学生有着多方面意义。一方面，矩阵有助于我们理解和学习各种数学算法，这些算法对 AI 意义重大，也有助于我们提高程序效率。另一方面，学习矩阵论有利于培养学生的逻辑运算思维，拓展数学视野，对生活中的事物有了更深刻的理解。

1.1.2 函数矩阵和矩阵函数介绍

很多实际问题可以转化为求解线性系统 Ax=b，其中 A=A(f1(D)),b=b(f2(D)),D 是一个 N 阶方阵。那么 f1(D)和 f2(D)分别是关于矩阵 D 的矩阵函数，矩阵 A 和向量 b 分别是关于 f1(D)和 f2(D)的函数矩阵。因此，我们需要进一步学习函数矩阵和矩阵函数的性质和求解方法。

1.矩阵函数

顾名思义，矩阵函数即以 n 阶矩阵为自变量和函数值(因变量)的一种函数，书中对矩阵函数的定义如下:

定义 3.7　设一元函数 $f(z)$ 能够展开为 z 的幂级数

$$f(z)=\sum_{k=0}^{\infty}c_k z^k \quad (|z|<r) \qquad (3.3.1)$$

其中，$r>0$ 表示该幂级数的收敛半径．当 n 阶矩阵 $\mathbf{A}$ 的谱半径 $\rho(\mathbf{A})<r$ 时，把收敛的矩阵幂级数 $\sum_{k=0}^{\infty}c_k\mathbf{A}^k$ 的和称为矩阵函数，记为 $f(\mathbf{A})$，即

$$f(\mathbf{A})=\sum_{k=0}^{\infty}c_k\mathbf{A}^k \qquad (3.3.2)$$

一些常见的矩阵函数包括：

$$e^z = \sum_{m=0}^{\infty} \frac{1}{m!} z^m, \quad R = +\infty,$$

$$\sin z = \sum_{m=0}^{\infty} \frac{(-1)^m}{(2m+1)!} z^{2m+1}, \quad R = +\infty,$$

$$\cos z = \sum_{m=0}^{\infty} \frac{(-1)^m}{(2m)!} z^{2m}, \quad R = +\infty,$$

$$(1-z)^{-1} = \sum_{m=0}^{\infty} z^m, \quad R = 1,$$

相应的有矩阵函数：

$$e^A = \sum_{m=0}^{\infty} \frac{1}{m!} A^m, \ \forall A \in \mathbb{C}^{n\times n},$$

$$\sin A = \sum_{m=0}^{\infty} \frac{(-1)^m}{(2m+1)!} A^{2m+1}, \ \forall A \in \mathbb{C}^{n\times n},$$

$$\cos A = \sum_{m=0}^{\infty} \frac{(-1)^m}{(2m)!} A^{2m}, \ \forall A \in \mathbb{C}^{n\times n},$$

$$(I-A)^{-1} = \sum_{m=0}^{\infty} A^m, \ \rho(A) < 1.$$

矩阵函数和普通函数一样具有很多性质，如：

命题 1 设 $A \in \mathbb{C}^{n\times n}$，则 1）$e^{iA} = \cos A + i\sin A$；

2）$\cos A = \frac{1}{2}(e^{iA} + e^{-iA})$；

3）$\sin A = \frac{1}{2i}(e^{iA} - e^{-iA})$；

4）$\sin^2 A + \cos^2 A = I$；

5）若 $AB = BA$，则 $e^A e^B = e^B e^A = e^{A+B}$；6）一般地，$e^A e^B$，$e^B e^A$，$e^{A+B}$ 互不相等；7）$e^A e^{-A} = e^{-A} e^A = I$，$(e^A)^{-1} = e^{-A}$.

矩阵函数的求法具有规律性，包括：待定系数法，数项级数求和法，对角型法，若尔当标准型法等，第三章将详细讨论这些求法。

2.函数矩阵

顾名思义，函数矩阵即矩阵中的每一元素a_{ij}都为一函数，包括单变量函数和多变量函数。函数矩阵的求导和积分是作用在各个矩阵元素上，没有更多的规则。

1.1.3 线性代数方程组求解介绍

含 m 个方程，n 个未知数的线性方程组的一般形式如下：

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases}$$

即 Ax=b。

（1）求解齐次线性方程组 Ax=0

由于齐次线性方程组一定有平凡解 x=0，所以没有解不存在的情况。

判定解的个数：

齐次线性方程组有唯一解的充分必要条件是 r (A) = n，即系数矩阵列满秩。

齐次线性方程组有基础解系（无穷解）的充分必要条件是 r (A) < n

求解：通过初等行变换将矩阵化为阶梯型矩阵=>重新写回方程式形式=>得到基础解系=>写出方程组通解

（2）求解非齐次线性方程组

判定解的个数:

非齐次线性方程组无解的充分必要条件是 r(A)<r(B)，即系数矩阵的秩小于增广矩阵的秩。

非齐次线性方程组有唯一解的充分必要条件是 r(A)=r(B)=n

非齐次线性方程组有无穷解的充分必要条件是 r(A)=r(B)<n

求解：将增广矩阵通过初等行变换化为阶梯型矩阵=>得到同解方程组=>导出基础解系=>求非齐次组的特解=>原方程组的通解=特解+通解

例：

1.2 问题介绍

1.2.1 矩阵函数的求法问题介绍

很多情况下，已知矩阵 A，并不能直接地计算出矩阵函数值，因此我们需要采用将矩阵分解等一些更加巧妙的方法。本文将主要介绍待定系数法、数项级数求和法、对角型法、若当标准形法。这些方法在生活中能够优化算法的复杂度，简化解题策略。

1.2.2 矩阵分解的方法问题介绍

矩阵分解是一种将矩阵简化为其组成部分的方法。这种方法可以简化更复杂的矩阵运算，这些运算可以在分解的矩阵上执行，而不是在原始矩阵本身上执行。寻求矩阵在各种意义下的分解形式，对矩阵相关的数值计算和理论有着深刻意义。不同的分解式，能够不同程度地反映出原矩阵的某些特征；同时，分解的方法与过程也提供了某些有效的数值计算方法，并为理论分析提供了依据。矩阵分解也有助于解决最优化问题和数值代数问题等。

1.3 国内外研究现状

1.3.1 矩阵函数的求法研究现状

当前主流的矩阵函数求法包括利用矩阵标准型求矩阵函数 f(A)(利用其与对角阵相似或与 Jordan 标准型相似)、利用最小多项式、本文即将介绍的四种方法、利用零化多项式求矩阵函数等，求法种类多样，需要使用者根据矩阵特点和算法需求选择合适的求法。

1.3.2 矩阵分解方法研究现状

国内外对于矩阵分解的方法层出不穷，时刻更迭，不同的矩阵分解方法适用于不同特征的矩阵。常见的矩阵分解算法有：LU 分解，可以说是最简单的矩阵分解算法；Cholesky 分解是 LU 分解的进阶，只适用于正定的对称阵；QR 分解将原矩阵分解成一

个正交矩阵 Q 和一个上三角矩阵 R 的乘积; EVD 分解和 SVD 分解的应用范围同样十分广泛。矩阵分解是一个值得研究的方向，在生活中有很多有效的应用。

1.4 本论文工作简述

1.4.1 本论文对上述问题研究简述

本文参考程云鹏《矩阵论》、张贤达的《矩阵分析及其应用》以及国内外其它相关资料，结合笔者课程中的理解，对几种重要的矩阵函数求法和矩阵分解方法进行了总结，同时也对方法所需的相关知识进行了铺垫。本文分为预备知识、矩阵函数的求法研究、矩阵分解方法研究、总结四个部分，详细介绍了待定系数法、数项级数求和法、对角型法、若当标准型法四种矩阵函数求法，及矩阵的 LU 分解、QR 分解、满秩分解、奇异值分解四种矩阵分解方法。

1.4.2 本论文创新点或特点简述

本论文语言较通俗易懂，每种方法都包含有知识铺垫、步骤和举例说明，便于读者理解。同时包含有预备知识部分，为相关方法进行了知识铺垫。

1.4.3 本论文撰写结构简述

本文分为预备知识、矩阵函数的求法研究、矩阵分解方法研究、总结四个部分，预备知识部分对欧氏空间与线性变换、向量范数、矩阵范数、矩阵函数展开了大致的介绍，第三、四章详细介绍了待定系数法、数项级数求和法、对角型法、若当标准型法四种矩阵函数求法，及矩阵的 LU 分解、QR 分解、满秩分解、奇异值分解四种矩阵分解方法。第五章对相关方法进行了总结。

第二章 预备知识

2.1 欧式空间与线性变换

2.1.1 欧式空间与线性变换介绍

1.欧式空间

（1）定义：若欧氏空间满足以下性质：

（1）**交换律**：$(x, y)=(y, x)$；

（2）**分配律**：$(x, y+z)=(x, y)+(x, z)$；

（3）**齐次性**：$(kx, y)=k(x, y)\ (\forall k \in \mathbf{R})$；

（4）**非负性**：$(x, x)\geqslant 0$，**当且仅当** $x=0$ **时**，$(x, x)=0$.

则称 V 为欧式空间。

（2）内积的性质

V为欧氏空间， $\forall x, y, z \in V, \forall k \in R$

1) $(x, ky)=k(x, y),\quad (kx, ky)=k^2(x, y)$

2) $(0, y)=(x, 0)=0$

3) $(x, y+z)=(x, y)+(x, z)$

（3）n 维欧氏空间中内积的矩阵表示

（4）欧氏空间中向量 x 的长度的长度记为：$|x|=\sqrt{(x, x)}$

（5）向量长度有以下性质：

1) $|x|\geq 0;\quad |x|=0 \Leftrightarrow x=0$

2) $|kx|=|k||x|$

3）非零向量 x **的单位化：** $\dfrac{1}{|x|}x$.

（6）若（x,y）=0，则 x 与 y 正交或垂直。

2.线性变换

（1）线性空间

线性空间是一个较为常见的代数系统。它的定义为：

定义 1.1 设 V 是一个非空集合，它的元素用 $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{z}$ 等表示，并称之为向量；K 是一个数域，它的元素用 k, l, m 等表示．如果 V 满足下列条件：

(1) 在 V 中定义一个加法运算，即当 $\boldsymbol{x}, \boldsymbol{y} \in V$ 时，有唯一的和 $\boldsymbol{x} + \boldsymbol{y} \in V$，且加法运算满足下列性质

1) 结合律 $\boldsymbol{x} + (\boldsymbol{y} + \boldsymbol{z}) = (\boldsymbol{x} + \boldsymbol{y}) + \boldsymbol{z}$；

2) 交换律 $\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{y} + \boldsymbol{x}$；

3) 存在零元素 $\boldsymbol{0}$，使 $\boldsymbol{x} + \boldsymbol{0} = \boldsymbol{x}$；

4) 存在负元素，即对任一向量 $\boldsymbol{x} \in V$，存在向量 $\boldsymbol{y} \in V$，使 $\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{0}$，则称 $\boldsymbol{y}$ 为 $\boldsymbol{x}$ 的负元素，记为 $-\boldsymbol{x}$，于是有

$$\boldsymbol{x} + (-\boldsymbol{x}) = \boldsymbol{0}$$

(2) 在 V 中定义数乘（数与向量的乘法）运算，即当 $\boldsymbol{x} \in V$，$k \in K$ 时，有唯一的 $k\boldsymbol{x} \in V$，且数乘运算满足下列性质：

5) 数因子分配律 $k(\boldsymbol{x} + \boldsymbol{y}) = k\boldsymbol{x} + k\boldsymbol{y}$；

6) 分配律 $(k + l)\boldsymbol{x} = k\boldsymbol{x} + l\boldsymbol{x}$；

7) 结合律 $k(l\boldsymbol{x}) = (kl)\boldsymbol{x}$；

8) $1\,\boldsymbol{x} = \boldsymbol{x}$.

则称 V 为数域 K 上的线性空间或向量空间．

（2）如果数域 K 上的线性空间 V 的变换 T 有性质：

$T(k\boldsymbol{x} + l\boldsymbol{y}) = k(T\boldsymbol{x}) + l(T\boldsymbol{y})$ 则称 T 为 V 的一个线性变换。

（3）线性变换的运算

在线性变换上定义了一系列运算。

和的定义如下：

设 T_1, T_2 为线性空间V的两个线性变换，定义它们的和 $T_1 + T_2$ 为：$(T_1 + T_2)(x) = T_1x + T_2x, \forall x \in V$

则 $T_1 + T_2$ 也是V的线性变换.

和的性质如下：

（1）满足交换律：$T_1 + T_2 = T_2 + T_1$

（2）满足结合律：$(T_1 + T_2) + T_3 = T_1 + (T_2 + T_3)$

（3） $T_0 + T_1 = T_1$ ，T_0 为零变换.

（4） $(-T) + T = T_0$

负变换的定义如下：

设 T 为线性空间V的线性变换，定义变换 $-T$ 为：

$$(-T)(x)=-T(x),\qquad \forall x\in V$$

则 $-T$ 也为V的线性变换，称之为 T 的负变换.

数量乘积的定义如下：

设 T 为线性空间V的线性变换，$k\in K$，定义 k 与 T 的数量乘积 kT 为：

$$(kT)(x)=kT(x),\quad \forall x\in V$$

则 kT 也是V的线性变换.

数量乘积的性质如下：

(1) $k(T_1+T_2)=kT_1+kT_2$

(2) $(k+l)T=kT+lT$

(3) $(kl)T=k(lT)$

(4) $1T=T$

乘积的定义如下： $(T_1T_2)(x)=T_1(T_2x),\quad \forall x\in V$

乘积的性质如下：

（1）满足结合律： $(T_1T_2)T_3=T_1(T_2T_3)$

（2） $T_eT=TT_e=T$ **，** T_e **为单位变换**

（3）交换律一般不成立，即一般地，

$$T_1T_2\neq T_2T_1$$

（4）乘法对加法满足左、右分配律：

$$T_1(T_2+T_3)=T_1T_2+T_1T_3$$

$$(T_1+T_2)T_3=T_1T_3+T_2T_3$$

线性变换的逆的定义如下：

设 T 为线性空间V的线性变换，若有V的变换 S 使

$$ST=TS=T_e$$

则称 T 为可逆变换，称 S 为 T 的逆变换，记作 T^{-1}.

线性变换的多项式的定义如下：

$$T^n=\underbrace{T\cdots T}_{n}$$

$$f(T)=a_mT^m+\cdots+a_1T+a_0T_e$$

线性变换的矩阵的定义如下：

设 $x_1,x_2,\cdots,x_n$为数域P上线性空间V的一组基，T为V的线性变换．基向量的象可以被基线性表出,设

$$\begin{cases} T(x_1)=a_{11}x_1+a_{21}x_2+\cdots+a_{n1}x_n \\ T(x_2)=a_{12}x_1+a_{22}x_2+\cdots+a_{n2}x_n \\ \qquad\cdots\cdots \\ T(x_n)=a_{1n}x_1+a_{2n}x_2+\cdots+a_{nn}x_n \end{cases}$$

2.1.2 若当标准型的求解

1.相关概念

Jordan 标准型的定义：

(1)形如 $J=\begin{bmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda \end{bmatrix}$ 的 m 阶方阵称为 m 阶 Jordan 块，其中 λ 可以是实数也可以是复数。以若干个 Jordan 块为对角元组成的块对角矩阵称为 Jordan 矩阵。

(2)不变因子：

可以证明[1]**，一个多项式矩阵 $\mathbf{A}(\lambda)$ 的标准形式(1.2.37)的对角线上的非零元素 $d_i(\lambda)(i=1,2,\cdots,s)$ 不随矩阵的初等变换而改变．因此，通常称 $d_i(\lambda)(i=1,2,\cdots,s)$ 为 $\mathbf{A}(\lambda)$ 的不变因子或不变因式．**

初等因子组：所有初等因子的集合。

2.求解若当标准型有以下步骤：

(1)求$\lambda E-A$的初等因子组

用初等变换将矩阵化为对角矩阵，若记对角线上的非零多项式为 $f_i(\lambda)$ ，那么诸次数大于零的 0 的 $f_i(\lambda)$ 的全体不可约因式，就是 $\mathbf{A}(\lambda)$ 的初等因子组。

(2)写出每个初等因子对应的 Jordan 块

$$\mathbf{J}_i(\lambda_i)=\begin{bmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \lambda_i & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_i \end{bmatrix}_{m_i\times m_i}$$

(3)写出以这些 Jordan 块构成的标准形

$$J=\begin{bmatrix} J_1(\lambda_1) & & & \\ & J_2(\lambda_2) & & \\ & & \ddots & \\ & & & J_s(\lambda_s) \end{bmatrix}$$

2.1.3 欧式空间中线性变换的求法(可参考课本例 1.36 和 ppt)

设 V 是欧氏空间，T 是 V 上的一个线性变换，求 $z=(T^k)(x)$，则则其步骤如下：

1. 任意找一组基，利用 Schmidt 正交化方法得到 V 的一组标准正交基 e1,…,en,x= $k_1e_1+...+k_ne_n$,其中 $k_i=(x,e_i)$.

求 T 在基下的矩阵 A_0,则 $T(e_1,...,e_n)=(e_1,...,e_n)A_0$

其中 $A_0=PJP^{-1}$，则 $T(e_1,...,e_n)=(e_1,...,e_n)PJP^{-1}$

那么 $T(e_1,...,e_n)P=(e_1,...,e_n)PJ$

2.得到一组新的基 $(E_1,...,E_n)=(e_1,...,e_n)P$

3.通过坐标变换得到 $x=(E_1,...,E_n)P^{-1}\begin{pmatrix}k_1\\ \vdots\\ k_n\end{pmatrix}=(E_1,...,E_n)\begin{pmatrix}l_1\\ \vdots\\ l_n\end{pmatrix}$

4.可以得到 T 在新基下的矩阵： $T(E_1,...,E_n)=(E_1,...,E_n)J$

5. $T(x)=(E_1,...,E_n)J\begin{pmatrix}l_1\\ \vdots\\ l_n\end{pmatrix}\Rightarrow(T^k)(x)=(E_1,...,E_n)J^k\begin{pmatrix}l_1\\ \vdots\\ l_n\end{pmatrix}$

2.2 向量范数与矩阵范数

2.2.1 向量范数介绍

1.向量范数的定义

定义 2.1 如果 V 是数域 K 上的线性空间，且对于 V 的任一向量 $\boldsymbol{x}$，对应一个实值函数 $\|\boldsymbol{x}\|$，它满足以下三个条件：

(1) 非负性：当 $\boldsymbol{x}\neq\boldsymbol{0}$ 时，$\|\boldsymbol{x}\|>0$；当 $\boldsymbol{x}=\boldsymbol{0}$ 时，$\|\boldsymbol{x}\|=0$；

(2) 齐次性：$\|a\boldsymbol{x}\|=|a|\,\|\boldsymbol{x}\|$ $(a\in K,\ \boldsymbol{x}\in V)$；

(3) 三角不等式：$\|\boldsymbol{x}+\boldsymbol{y}\|\leqslant\|\boldsymbol{x}\|+\|\boldsymbol{y}\|$ $(\boldsymbol{x},\ \boldsymbol{y}\in V)$.

则称 $\|\boldsymbol{x}\|$ 为 V 上向量 $\boldsymbol{x}$ 的范数，简称**向量范数**.

2.向量范数的性质

(1) 当 $\|x\|\neq 0$ 时，$\left\|\frac{1}{\|x\|}x\right\|=1$ $\quad\because \left\|\frac{1}{\|x\|}x\right\|=\frac{1}{\|x\|}\|x\|=1$

(2) $\forall x\in V$ ，$\|-x\|=\|x\|$ $\quad\because \|-x\|=|-1|\|x\|=\|x\|$

(3) $\forall x,y\in V$ ，$\|x\|-\|y\|\leq\|x-y\|$

$\because \|x\|=\|(x-y)+y\|\leq\|x-y\|+\|y\|\Rightarrow\|x\|-\|y\|\leq\|x-y\|$

(4) $\forall x,y\in V$ ，$\left|\|x\|-\|y\|\right|\leq\|x-y\|$

$\|x\|=\|(x-y)+y\|\leq\|x-y\|+\|y\| \Rightarrow\|x\|-\|y\|\leq\|x-y\|$

同样 $\|y\|-\|x\|\leq\|y-x\|=\|x-y\|$

3.以下为一些常见的向量范数：

1：$\|x\|_1=\sum|\xi_i|$ 是一种向量范数，记为**1-范数**

2：$\|x\|=\sqrt{(x,x)}$ 是一种向量范数，记**2-范数**

3: $\|x\|=\max_i|x_i|$ 是一种向量范数,记为 ∞ -范数

4.向量范数之间存在等价性：

定理 2.1 设 $\|x\|_\alpha$ 和 $\|x\|_\beta$ 为有限维线性空间 V 的任意两种向量范数(它们不限于 p-范数)，则存在两个与向量 x 无关的正常数 c_1 和 c_2，使得不等式

$$c_1\|x\|_\beta\leqslant\|x\|_\alpha\leqslant c_2\|x\|_\beta \quad (\forall x\in V) \quad (2.1.9)$$

成立.

2.2.2 矩阵范数介绍

1.矩阵范数的概念

对向量范数进行推广，就获得了矩阵范数，因此矩阵范数也是对矩阵的一种度量方式。它满足以下条件：

(1) 非负性：当 $A\neq O$ 时，$\|A\|>0$；当 $A=O$ 时，$\|A\|=0$；

(2) 齐次性：$\|\alpha A\|=|\alpha|\,\|A\|$ $(\alpha\in\mathbf{C})$；

(3) 三角不等式：$\|A+B\|\leqslant\|A\|+\|B\|$ $(B\in\mathbf{C}^{m\times n})$.

(4) 相容性：

$$\|AB\|\leqslant\|A\|\,\|B\| \quad (B\in\mathbf{C}^{n\times l}) \quad (2.2.1)$$

相容性：

定义 2.4 对于 $\mathbf{C}^{m\times n}$ 上的矩阵范数 $\|\cdot\|_M$ 和 $\mathbf{C}^m$ 与 $\mathbf{C}^n$ 上的同类向量范数 $\|\cdot\|_V$，如果

$$\|\boldsymbol{Ax}\|_V \leqslant \|\mathbf{A}\|_M \|\boldsymbol{x}\|_V \quad (\forall \mathbf{A} \in \mathbf{C}^{m\times n},\ \forall \boldsymbol{x} \in \mathbf{C}^n) \tag{2.2.2}$$

则称矩阵范数 $\|\cdot\|_M$ 与向量范数 $\|\cdot\|_V$ 是相容的．

矩阵的 Frobenius 范数为：

$$\|A\|_{m2} = \left[\mathbf{tr}\left(A^{\mathrm{H}}A\right)\right]^{\frac{1}{2}} = \left[\mathbf{tr}\left(AA^{\mathrm{H}}\right)\right]^{\frac{1}{2}}$$ 称为矩阵的Frobenius范数，记做 $\|A\|_{\mathrm{F}}$

矩阵范数具有等价性。

2.从属范数

定理：对 C^m与 C^n上的同类向量范数 $\|x\|_V$，定义

$$\|A\| = \max_{\|x\|_V=1} \|Ax\|_V \qquad \left(\forall A_{m\times n}, x \in C^n\right)$$

则 $\|A\|$ 是 $C^{m\times n}$ 中矩阵A 的范数，且 $\|A\|$ 与 $\|x\|_V$ 相容

$\|A\|$称为由 $\|x\|_V$导出的矩阵范数（或称为从属范数）

列和范数：$$\|A\|_1 = \max_j \left\{\sum_{i=1}^{m} |a_{ij}|\right\}$$

谱范数：$$\|A\|_2 = \sqrt{\lambda_1}, \lambda_1 = \max\left\{\lambda\left(A^H A\right)\right\}$$

行和范数：$$\|A\|_\infty = \max_i \left\{\sum_{j=1}^{n} |a_{ij}|\right\}$$

2.2.3 矩阵可逆性条件、谱半径和条件数介绍

1. 矩阵可逆性条件

若存在 A 的某范数，使得 ‖A‖<1，则 I – A 可逆。

假设 不可逆，则$(I - A)x = 0$存在非零解.

2. 谱半径

矩阵的谱半径在在特征值估计、广义逆矩阵、数值分析以及数值代数等理论的建树中都占有极其重要的地位。

（1）定义：

定义 2.5 设 $\mathbf{A} \in \mathbf{C}^{n\times n}$ 的 n 个特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$，称

$$\rho(\mathbf{A}) = \max_i |\lambda_i| \tag{2.3.3}$$

为 $\mathbf{A}$ 的谱半径．

（2）性质：

$$\forall A \in C^{n\times n}, \forall \|\bullet\|_M，\ 有\ \ \rho(A) \le \|A\|_M$$

定理 2.10 设 $\mathbf{A} \in \mathbf{C}^{n\times n}$,对任意的正数 ε,存在某种矩阵范数 $\|\cdot\|_M$,使得

$$\|\mathbf{A}\|_M \leqslant \rho(\mathbf{A}) + \varepsilon \qquad (2.3.7)$$

3.条件数:

令 $\mathrm{cond}(\mathbf{A}) = \|\mathbf{A}\|\,\|\mathbf{A}^{-1}\|$,则称 cond(A)为矩阵 A 的条件数,它是求矩阵逆的摄动的一个重要量.一般说来条件数愈大, $(\mathbf{A}+\delta\mathbf{A})^{-1}$ 与 $\mathbf{A}^{-1}$ 的相对误差就愈大.

2.3 矩阵函数介绍

2.3.1 矩阵序列介绍

1.矩阵序列的收敛和发散定义如下:

定义:将矩阵序列 $A^{(k)} = \left(a_{ij}^{(k)}\right)_{m\times n}$,记作 $\{A^{(k)}\}$

当 $\lim\limits_{k\to\infty} a_{ij}^{(k)} = a_{ij}(\forall i,j)$ 时,称矩阵序列 $\{A^{(k)}\}$ 收敛于矩阵 $A=(a_{ij})$。记作

$$\lim_{k\to\infty} A^{(k)} = A \text{ ,或者 } A^{(k)} \to A(k\to\infty)$$

若数列 $\left(a_{ii}^{(k)}\right)$ 之一发散,称 $\{A^{(k)}\}$ 发散

2.性质

(1) 若 $\lim\limits_{k\to\infty} A^{(k)} = A_{m\times n}, \lim\limits_{k\to\infty} B^{(k)} = B_{m\times n}$ 则

$$\lim_{k\to\infty}(aA^{(k)} + bB^{(k)}) = aA + bB, \quad \forall a,b$$

(2) 若 $\lim\limits_{k\to\infty} A^{(k)} = A_{m\times n}, \lim\limits_{k\to\infty} B^{(k)} = B_{n\times l}$ 则

$$\lim_{k\to\infty}(A^{(k)}B^{(k)}) = AB$$

(3) 若 $A^{(k)}$ 与 A 是可逆矩阵,且 $\lim\limits_{k\to\infty} A^{(k)} = A$,则

$$\lim(A^{(k)})^{-1} = A^{-1}$$

3.定理

定理1:设 $A^{(k)}, A \in C^{m\times n}$,则

(1) $\lim\limits_{k\to\infty} A^{(k)} = 0 \iff \forall \|\bullet\|, \lim\limits_{k\to\infty} \left\|A^{(k)}\right\| = 0$

(2) $\lim\limits_{k\to\infty} A^{(k)} = A \iff \forall \|\bullet\|, \lim\limits_{k\to\infty} \|A^{(k)} - A\| = 0$

定理2: A为收敛矩阵 $\iff \rho(A) < 1$

定理3:若矩阵范数 $\|\bullet\|_M$使 $\|A\|_M < 1$,则 $A^k \to 0$

2.3.2 矩阵级数介绍

1.定义

定义 3.4 把定义 3.1 中的矩阵序列所形成的无穷和 $\boldsymbol{A}^{(0)}+\boldsymbol{A}^{(1)}+\boldsymbol{A}^{(2)}+\cdots+\boldsymbol{A}^{(k)}+\cdots$ 称为矩阵级数，记为 $\sum\limits_{k=0}^{\infty}\boldsymbol{A}^{(k)}$，即有

$$\sum_{k=0}^{\infty}\boldsymbol{A}^{(k)}=\boldsymbol{A}^{(0)}+\boldsymbol{A}^{(1)}+\boldsymbol{A}^{(2)}+\cdots+\boldsymbol{A}^{(k)}+\cdots \quad (3.2.1)$$

2.敛散性

若 $\lim\limits_{N\to\infty}\boldsymbol{S}^{(N)}=\boldsymbol{S}$，称 $\sum\limits_{k=0}^{\infty}A^{(k)}$ 收敛于$\boldsymbol{S}$，记做

$$\sum_{k=0}^{\infty}A^{(k)}=\boldsymbol{S}$$

若 $\{\boldsymbol{S}^{(N)}\}$ 发散，称 $\sum\limits_{k=0}^{\infty}A^{(k)}$ 发散

3.性质

性质1：$\sum\limits_{k=0}^{\infty}A^{(k)}=S \Longleftrightarrow \sum\limits_{k=0}^{\infty}a_{ij}^{(k)}=s_{ij}\ (all\quad i,j)$

性质2：若 $\sum\limits_{k=0}^{\infty}\left|a_{ij}^{(k)}\right|$收敛 $(all\quad i,j)$，称 $\sum\limits_{k=0}^{\infty}A^{(k)}$ 绝对收敛。

（1）$\sum\limits_{k=0}^{\infty}A^{(k)}$绝对收敛 $\Rightarrow\sum\limits_{k=0}^{\infty}A^{(k)}$ 收敛

（2）若$\sum\limits_{k=0}^{\infty}A^{(k)}$绝对收敛于$\boldsymbol{S}$，对$\sum\limits_{k=0}^{\infty}A^{(k)}$ 任意重组重排得 $\sum\limits_{k=0}^{\infty}B^{(k)}$，则 $\sum\limits_{k=0}^{\infty}B^{(k)}$ 绝对收敛于$\boldsymbol{S}$。

性质3：$\sum\limits_{k=0}^{\infty}A^{(k)}=S$ 绝对收敛 $\Leftrightarrow\forall\|\cdot\|,\sum\limits_{k=0}^{\infty}\left\|A^{(k)}\right\|$ 收敛

性质4：$\sum\limits_{k=0}^{\infty}A^{(k)}$ 收敛于$S\Rightarrow\sum\limits_{k=0}^{\infty}PA^{(k)}Q$ 收敛于 PSQ

$\sum\limits_{k=0}^{\infty}A^{(k)}$ 绝对收敛$\Rightarrow\sum\limits_{k=0}^{\infty}PA^{(k)}Q$ 绝对收敛

性质5：$\sum\limits_{k=0}^{\infty}A^{(k)}$ 绝对收敛于$S_{m\times n}$,$\sum\limits_{k=0}^{\infty}B^{(k)}$绝对收敛于 $T_{n\times l}$

4.矩阵幂级数

对函数 $f(z)=\sum c_k z^k,(|z|<r)$，方阵 $A_{n\times n}$，构造矩阵幂级数 $f(A)=\sum c_k A^k$，

若 $\rho(A)<r$，则$\sum c_k A^k$绝对收敛；若 $\rho(A)>r$，$\sum c_k A^k$ 发散。

2.3.3 矩阵函数介绍

矩阵函数即以矩阵为自变量且取值为矩阵的函数。其定义为：

定义：设一元函数 $f(z)$ 能展开为z的幂级数

$$f(z)=\sum_{k=0}^{\infty}c_k z^k \quad \left(|z|<r, r>0\right)$$

其中r>0表示该幂级数的收敛半径。当n阶矩阵A的谱半径$\rho(A)<r$ 时，把收敛的矩阵幂级数$\sum_{k=0}^{\infty}c_k A^k$ 的和为f(A), 即 $f(A)=\sum_{k=0}^{\infty}c_k A^k$

在矩阵函数中，$e^A e^B = e^B e^A = e^{A+B}$一般不成立。（除 AB=BA 的情况）

2.4 函数矩阵对矩阵的导数

1.定义

定义: 如果矩阵 $A(t)=\left(a_{ij}(t)\right)_{m\times n}$,的每一个元素 $a_{ij}(t)$ 是变量t 的可微函数，则$A(t)$关于t的导数(微商)定义为

$$\frac{dA(t)}{dt}=\left(a'_{ij}(t)\right)_{m\times n}，\quad 或者 \quad A'(t)=\left(a'_{ij}(t)\right)_{m\times n}$$

2.定理

定理8：设 $A(t), B(t)$ 可导，则有

（1） $\frac{d}{dt}\left[A(t)+B(t)\right]=\frac{d}{dt}A(t)+\frac{d}{dt}B(t)$

（2） $A_{m\times n}, f(t)$ 可导 $\frac{d}{dt}\left[f(t)A(t)\right]=f'(t)A(t)+f(t)A'(t)$

（3） $A_{m\times n}, A_{n\times l}: \frac{d}{dt}\left[A(t)B(t)\right]=A'(t)B(t)+A(t)B'(t)$

定理9：设 $A_{n\times n}$ 为数量矩阵，则有

（1） $\frac{d}{dt}e^{tA}=Ae^{tA}=e^{tA}A$

（2） $\frac{d}{dt}\cos(tA)=-A\bullet\sin(tA)=-\sin(tA)\bullet A$

（3） $\frac{d}{dt}\sin(tA)=A\bullet\cos(tA)=\cos(tA)\bullet A$

第三章 矩阵函数的求法研究

3.1 待定系数法

3.1.1 待定系数法求矩阵函数的步骤推导

1.引入

设 n 阶矩阵 $\mathbf{A}$ 的特征多项式为 $\varphi(\lambda)=\det(\lambda \boldsymbol{I}-\mathbf{A})$. 如果首 1 多项式

$$\psi(\lambda)=\lambda^m+b_1\lambda^{m-1}+\cdots+b_{m-1}\lambda+b_m \quad (1\leqslant m\leqslant n) \tag{3.3.9}$$

满足:① $\psi(\mathbf{A})=\boldsymbol{O}$;② $\psi(\lambda)$ 整除 $\varphi(\lambda)$(矩阵 $\mathbf{A}$ 的最小多项式与特征多项式均满足这些条件). 那么,$\psi(\lambda)$ 的零点都是 $\mathbf{A}$ 的特征值. 记 $\psi(\lambda)$ 的互异零点为 $\lambda_1,\lambda_2,\cdots,\lambda_s$,相应的重数为 $r_1,r_2,\cdots,r_s$ $(r_1+r_2\cdots+r_s=m)$,则有

$$\psi^{(l)}(\lambda_i)=0 \quad (l=0,1,\cdots,r_i-1;\ i=1,2,\cdots,s)$$

这里,$\psi^{(l)}(\lambda)$ 表示 $\psi(\lambda)$ 的 l 阶导数(下同). 设 $f(z)=\sum_{k=0}^{\infty}c_kz^k=\psi(z)g(z)+r(z)$,其中 $r(z)$ 是次数低于 m 的多项式,于是可由

$$f^{(l)}(\lambda_i)=r^{(l)}(\lambda_i) \quad (l=0,1,\cdots,r_i-1;\ i=1,2,\cdots,s) \tag{3.3.11}$$

确定出 $r(z)$. 利用 $\psi(\mathbf{A})=\boldsymbol{O}$,可得

$$f(\mathbf{A})=\sum_{k=0}^{\infty}c_k\mathbf{A}^k=r(\mathbf{A}) \tag{3.3.12}$$

2.前提

如果首1多项式 $\psi(\lambda)=\lambda^m+a_1\lambda^{m-1}+\ldots+a_{m-1}\lambda+a_m$ $(1\le m\le n)$

满足 $\psi(\lambda)|\varphi(\lambda)$，分解

$$\psi(\lambda)=(\lambda-\lambda_1)^{m_1}\cdots(\lambda-\lambda_s)^{m_s} \quad \lambda_i\ne\lambda_j,\sum m_i=m$$

3.步骤

（1）将求 f（A）转化为求 r（A）

设 A 的特征多项式为 $\varphi(\lambda)=\det(\lambda\boldsymbol{I}-\mathbf{A})$，若首 1 多项式 $\psi(\lambda)=\lambda^m+b_1\lambda^{m-1}+\cdots+b_{m-1}\lambda+b_m$，满足① $\psi(\mathbf{A})=\boldsymbol{O}$;② $\psi(\lambda)$ 整除 $\varphi(\lambda)$，则有 $\psi^{(l)}(\lambda_i)=0$.设 $f(z)=\sum_{k=0}^{\infty}c_kz^k=\psi(z)g(z)+r(z)$, r(z)是次数低于 m 次的多项式。

（2）通过求导，带入特征值，解方程求出 r(A)

$$f^{(l)}(\lambda_i)=r^{(l)}(\lambda_i) \quad (l=0,1,\cdots,r_i-1;\ i=1,2,\cdots,s) \tag{3.3.11}$$

确定出 $r(z)$. 利用 $\psi(\mathbf{A})=\boldsymbol{O}$,可得

$$f(\mathbf{A})=\sum_{k=0}^{\infty}c_k\mathbf{A}^k=r(\mathbf{A}) \tag{3.3.12}$$

3.1.2 举例展示求法

设 $\mathbf{A}=\begin{bmatrix}2 & 0 & 0\\ 1 & 1 & 1\\ 1 & -1 & 3\end{bmatrix}$，求 $e^{\mathbf{A}}$

(1) 将 $f(A)$ 转化为求 $r(A)$.

$\varphi(\lambda)=det(\lambda I-A)=(\lambda-2)^3$. 易得 $m(\lambda)=(\lambda-2)^2$. 取 $\psi(\lambda)=(\lambda-2)^2$.

(2) 求导，求特征值，解方程求出 $r(A)$.

取 $f(\lambda)=e^{\lambda}$, $f(\lambda)=\psi(\lambda)g(\lambda)+(a+b\lambda)$, 则有

$$\begin{cases}f(2)=e^2\\ f'(2)=e^2\end{cases} \text{或} \begin{cases}a+2b=e^2\\ b=e^2.\end{cases}$$

解得 $a=-e^2$, $b=e^2$. $\therefore r(\lambda)=e^2(\lambda-1)$. 从而

$$e^A=f(A)=r(A)=e^2(A-I)=e^2\begin{bmatrix}1 & 0 & 0\\ 1 & 0 & 1\\ 1 & -1 & 2\end{bmatrix}.$$

3.2 数项级数求和法

3.2.1 数项级数求和法求矩阵函数的步骤推导

1. 利用首一多项式求出某一项与前一项或前几项的关系

设首 1 多项式 $\psi(\lambda)$ 形如式(3.3.9)，且满足 $\psi(\mathbf{A})=\mathbf{O}$，即

$$\mathbf{A}^m+b_1\mathbf{A}^{m-1}+\cdots+b_{m-1}\mathbf{A}+b_m\mathbf{I}=\mathbf{O}$$

或者

$$\mathbf{A}^m=k_0\mathbf{I}+k_1\mathbf{A}+\cdots+k_{m-1}\mathbf{A}^{m-1}\quad(k_i=-b_{m-i})$$

2. 利用关系化简矩阵函数

$$\mathbf{A}^m=k_0\mathbf{I}+k_1\mathbf{A}+\cdots+k_{m-1}\mathbf{A}^{m-1}\quad(k_i=-b_{m-i})$$

$$\begin{aligned}f(\mathbf{A})=\sum_{k=0}^{\infty}c_k\mathbf{A}^k&=(c_0\mathbf{I}+c_1\mathbf{A}+\cdots+c_{m-1}\mathbf{A}^{m-1})+\\&c_m(k_0\mathbf{I}+k_1\mathbf{A}+\cdots+k_{m-1}\mathbf{A}^{m-1})+\cdots+\\&c_{m+l}(k_0^{(l)}\mathbf{I}+k_1^{(l)}\mathbf{A}+\cdots+k_{m-1}^{(l)}\mathbf{A}^{m-1})+\cdots=\\&\left(c_0+\sum_{l=0}^{\infty}c_{m+l}k_0^{(l)}\right)\mathbf{I}+\left(c_1+\sum_{l=0}^{\infty}c_{m+l}k_1^{(l)}\right)\mathbf{A}+\cdots+\\&\left(c_{m-1}+\sum_{l=0}^{\infty}c_{m+l}k_{m-1}^{(l)}\right)\mathbf{A}^{m-1}\end{aligned}\quad(3.3.14)$$

3.2.2 举例展示求法

设 $A = \begin{bmatrix} \pi & 0 & 0 & 0 \\ 0 & -\pi & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$，求 $\sin A$.

$\varphi(\lambda) = \det(\lambda I - A) = \lambda^4 - \pi^2\lambda^2$. $\because \varphi(A) = 0$, 所以 $A^4 = \pi^2 A^2$, $A^5 = \pi^2 A^3$, $A^7 = \pi^4 A^3, \cdots$ 于是有 $\sin A = A - \frac{1}{3!}A^3 + \frac{1}{5!}A^5 - \frac{1}{7!}A^7 + \cdots$

$$= A - \frac{1}{3!}A^3 + \frac{1}{5!}\pi^2 A^3 - \frac{1}{7!}\pi^4 A^3 + \cdots$$

$$= A + \left(-\frac{1}{3!} + \frac{1}{5!}\pi^2 - \frac{1}{7!}\pi^4 + \frac{1}{9!}\pi^6 - \cdots\right)A^3$$

$$= A + \frac{\sin\pi - \pi}{\pi^3}A^3 = A - \pi^{-2}A^3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

3.3 对角型法

3.3.1 对角型法求矩阵函数的步骤推导

1.引入

当 A 与对角矩阵相似时，即有可逆矩阵 P，使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$。

则有 $\boldsymbol{A} = \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1}$，$\boldsymbol{A}^2 = \boldsymbol{P}\boldsymbol{\Lambda}^2\boldsymbol{P}^{-1}, \cdots$，于是可得

$$\sum_{k=0}^{N} c_k \boldsymbol{A}^k = \sum_{k=0}^{N} c_k \boldsymbol{P}\boldsymbol{\Lambda}^k\boldsymbol{P}^{-1} = \boldsymbol{P} \cdot \sum_{k=0}^{N} c_k \boldsymbol{\Lambda}^k \boldsymbol{P}^{-1} =$$

$$\boldsymbol{P}\begin{bmatrix} \sum_{k=0}^{N} c_k\lambda_1^k & & \\ & \ddots & \\ & & \sum_{k=0}^{N} c_k\lambda_n^k \end{bmatrix}\boldsymbol{P}^{-1}$$

从而

$$f(\boldsymbol{A})=\sum_{k=0}^{\infty}c_k\boldsymbol{A}^k=\boldsymbol{P}\begin{bmatrix}\sum\limits_{k=0}^{\infty}c_k\lambda_1^k & & \\ & \ddots & \\ & & \sum\limits_{k=0}^{\infty}c_k\lambda_n^k\end{bmatrix}\boldsymbol{P}^{-1}=$$

$$\boldsymbol{P}\begin{bmatrix}f(\lambda_1) & & \\ & \ddots & \\ & & f(\lambda_n)\end{bmatrix}\boldsymbol{P}^{-1}$$

2.步骤

（1）求出相似对角矩阵

（2）用相似矩阵替代 A，带入函数

3.3.2 举例展示求法

设 $\boldsymbol{A}=\begin{bmatrix}4 & 6 & 0\\ -3 & -5 & 0\\ -3 & -6 & 1\end{bmatrix}$**，求** e^{A}

1、求出相似对角矩阵

$\varphi(\lambda)=det(\lambda I-A)=(\lambda+2)(\lambda-1)^2$.

对应 $\lambda_1=-2$ 的特征向量 $p_1=(-1,1,1)^T$；对应 $\lambda_2=\lambda_3=1$ 的两个线性无关的特征向量 $p_2=(-2,1,0)^T$，$p_3=(0,0,1)^T$. 构造矩阵 $P=(p_1,p_2,p_3)=\begin{bmatrix}-1 & -2 & 0\\ 1 & 1 & 0\\ 1 & 0 & 1\end{bmatrix}$

则有 $P^{-1}=\begin{bmatrix}1 & 2 & 0\\ -1 & -1 & 0\\ -1 & -2 & 1\end{bmatrix}$. $P^{-1}AP=\begin{bmatrix}-2 & & \\ & 1 & \\ & & 1\end{bmatrix}$

2. 用相似矩阵替代A，带入函数.

$$e^{A}=P\begin{bmatrix}e^{-2} & & \\ & e & \\ & & e\end{bmatrix}P^{-1}=\begin{bmatrix}2e-e^{-2} & 2e-2e^{-2} & 0\\ e^{-2}-e & 2e^{-2}-e & 0\\ e^{-2}-e & 2e^{-2}-2e & e\end{bmatrix}.$$

3.4 若尔当标准型法

3.4.1 若尔当标准型求矩阵函数的步骤推导

1.引入

（1）Jordan 分解定理

定理 1.29 设 $\boldsymbol{A}$ 是 n 阶复矩阵，且其特征多项式的某种分解式是(1.2.32)，则存在 n 阶复非奇异矩阵 $\boldsymbol{P}$，使

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{J} \qquad (1.2.34)$$

（2）推导

设 $\boldsymbol{A}$ 的 Jordan 标准形为 $\boldsymbol{J}$，则有可逆矩阵 $\boldsymbol{P}$，使得

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{J}=\begin{bmatrix}\boldsymbol{J}_1 & & \\ & \ddots & \\ & & \boldsymbol{J}_s\end{bmatrix}$$

其中

$$\boldsymbol{J}_i=\begin{bmatrix}\lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \lambda_i & 1 \\ & & & \lambda_i\end{bmatrix}_{m_i\times m_i}$$

$$f(\boldsymbol{A})=\sum_{k=0}^{\infty}c_k\boldsymbol{A}^k=\sum_{k=0}^{\infty}c_k\boldsymbol{P}\boldsymbol{J}^k\boldsymbol{P}^{-1}=\boldsymbol{P}\left(\sum_{k=0}^{\infty}c_k\boldsymbol{J}^k\right)\boldsymbol{P}^{-1}=$$

$$\boldsymbol{P}\begin{bmatrix}\sum_{k=0}^{\infty}c_k\boldsymbol{J}_1^k & & \\ & \ddots & \\ & & \sum_{k=0}^{\infty}c_k\boldsymbol{J}_s^k\end{bmatrix}\boldsymbol{P}^{-1}=$$

$$\boldsymbol{P}\begin{bmatrix}f(\boldsymbol{J}_1) & & \\ & \ddots & \\ & & f(\boldsymbol{J}_s)\end{bmatrix}\boldsymbol{P}^{-1} \qquad (3.3.17)$$

可求得

$$f(\boldsymbol{J}_i)=\sum_{k=0}^{\infty}c_k\boldsymbol{J}_i^k=\sum_{k=0}^{\infty}c_k\begin{bmatrix}\lambda_i^k & \mathrm{C}_k^1\lambda_i^{k-1} & \cdots & \mathrm{C}_k^{m_i-1}\lambda_i^{k-m_i+1} \\ & \lambda_i^k & \ddots & \vdots \\ & & \ddots & \mathrm{C}_k^1\lambda_i^{k-1} \\ & & & \lambda_i^k\end{bmatrix}=$$

$$\begin{bmatrix}f(\lambda_i) & \frac{1}{1!}f'(\lambda_i) & \cdots & \frac{1}{(m_i-1)!}f^{(m_i-1)}(\lambda_i) \\ & f(\lambda_i) & \ddots & \vdots \\ & & \ddots & \frac{1}{1!}f'(\lambda_i) \\ & & & f(\lambda_i)\end{bmatrix}$$

这表明，矩阵幂级数的求和问题可以转化为求矩阵的 Jordan 标准形及变换矩阵的问题。

（3）步骤

a) 将 A 的矩阵函数转化为 J 的矩阵函数

b) 再转化为 Ji 的矩阵函数

c) 计算 Ji 的 k 次幂

d) 计算 f（Ji）

3.4.2 举例展示求法

设 $A=\begin{bmatrix}\pi & 0 & 0 & 0\\ 0 & -\pi & 0 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 0 & 0\end{bmatrix}$，求 $\sin A$.

A为Jordan标准形，它的三个Jordan块为

$J=\pi$，$J_2=-\pi$，$J_3=\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}$

$\sin J_1=\sin\pi=0$ $\sin J_2=\sin(-\pi)=0$ $\sin J_3=\begin{bmatrix}\sin 0 & 1\cdot\cos 0\\ 0 & \sin 0\end{bmatrix}=\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}$

$$\sin A=\begin{bmatrix}\sin J_1 & & \\ & \sin J_2 & \\ & & \sin J_3\end{bmatrix}=\begin{bmatrix}0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 0 & 0\end{bmatrix}$$

第四章 矩阵分解方法研究

4.1 矩阵的 LU 分解

4.1.1 矩阵 LU 分解的步骤推导

1.定义

L 指下三角矩阵，U 指上三角矩阵。矩阵的 LU 分解源于线性方程组的高斯消元过程。 对于一个含有 N 个变量的 N 个线性方程组，总可以用高斯消去法，把左边的系数矩阵分解为一个单位下三角矩阵和一个上三角矩阵。矩阵的 LU 分解即将一个矩阵分解为下三角矩阵和上三角矩阵的乘积，下三角阵对角元是 1，上三角是主元。

对于任意矩阵 A 可以写为：A=LU 或 A=LDU（D 为对角阵）.

2.引入

（1）A 左乘 E，即是对 A 作相应的初等行变换.若用 Gauss 消去法将矩阵 A 转化成一个阶梯形矩阵 U，相应的初等变换对应的矩阵为 E1, Λ ,E r 则 Er Λ E1A=U.

（2）

定义 4.2 设矩阵 $\boldsymbol{A}$ 有唯一的 LDU 分解．若把 $\boldsymbol{A}=\boldsymbol{LDU}$ 中的 $\boldsymbol{D}$ 与 $\boldsymbol{U}$ 结合起来，并且用 $\hat{\boldsymbol{U}}$ 来表示，就得到唯一的分解

$$\boldsymbol{A}=\boldsymbol{L}(\boldsymbol{DU})=\boldsymbol{L}\hat{\boldsymbol{U}} \tag{4.1.23}$$

称为 $\boldsymbol{A}$ 的 **Doolittle 分解**；若把 $\boldsymbol{A}=\boldsymbol{LDU}$ 中的 $\boldsymbol{L}$ 与 $\boldsymbol{D}$ 结合起来，并且用 $\hat{\boldsymbol{L}}$ 来表示，就得到唯一的分解

$$\boldsymbol{A}=(\boldsymbol{LD})\boldsymbol{U}=\hat{\boldsymbol{L}}\boldsymbol{U} \tag{4.1.24}$$

称为 $\boldsymbol{A}$ 的 **Crout 分解**．

（3）

定理 4.1 设 $\boldsymbol{A}=(a_{ij})$ 是 n 阶矩阵，则当且仅当 $\boldsymbol{A}$ 的顺序主子式 $\Delta_k \neq 0\ (k=1,2,\cdots,n-1)$ 时，$\boldsymbol{A}$ 可唯一地分解为 $\boldsymbol{A}=\boldsymbol{LDU}$，其中 $\boldsymbol{L}$ 是单位下三角矩阵，$\boldsymbol{U}$ 是单位上三角矩阵，$\boldsymbol{D}$ 是对角矩阵

$$\boldsymbol{D}=\operatorname{diag}(d_1, d_2, \cdots, d_n)$$

其中，$d_k=\dfrac{\Delta_k}{\Delta_{k-1}}\ (k=1,2,\cdots,n;\ \Delta_0=1)$.

（4）

定义 4.3 称式(4.1.32)为实对称正定矩阵的 Cholesky 分解（平方根分解、对称三角分解）.

3.步骤

若将 $L_{n-1}L_{n-2}\cdots L_1A=U$ 化为 LU 分解，步骤如下：

1.令 $L=L_1^{-1}\cdots L_{n-2}^{-1}L_{n-1}^{-1}$，则 A=LU

2.计算 Luc=b.

3.解 $Ly=b$，$Ux=y$

4.1.2 举例展示求法

以 3*3 的矩阵 A 为例.

$$A=\begin{bmatrix}1&2&3\\2&5&7\\3&5&3\end{bmatrix}.$$

先将第一列元素中 a_{11} 以下的所有元素变为0,即

$$L_1A=\begin{bmatrix}1&0&0\\-2&1&0\\-3&0&1\end{bmatrix}\times\begin{bmatrix}1&2&3\\2&5&7\\3&5&3\end{bmatrix}=\begin{bmatrix}1&2&3\\0&1&1\\0&-1&-6\end{bmatrix}$$

再将矩阵第二列元素中 a_{22} 以下的所有元素变为0,即

$$L_2(L_1A)=\begin{bmatrix}1&0&0\\0&1&0\\0&1&1\end{bmatrix}\times\begin{bmatrix}1&2&3\\0&1&1\\0&-1&-6\end{bmatrix}=\begin{bmatrix}1&2&3\\0&1&1\\0&0&-5\end{bmatrix}=U$$

$$L=L_1^{-1}L_2^{-1}=\begin{bmatrix}1&0&0\\2&1&0\\3&0&1\end{bmatrix}\times\begin{bmatrix}1&0&0\\0&1&0\\0&-1&1\end{bmatrix}=\begin{bmatrix}1&0&0\\2&1&0\\3&-1&1\end{bmatrix}.$$

4.2 矩阵的 QR 分解

4.2.1 矩阵的 QR 分解的步骤推导

1.定义

定义 4.6 如果实（复）非奇异矩阵 A 能够化成正交（酉）矩阵 Q 与实（复）非奇异上三角矩阵 R 的乘积，即

$$A = QR \tag{4.2.7}$$

则称式（4.2.7）为 A 的 QR 分解．

2.引入

（1）

定理 4.6 设 A 是 n 阶实（复）非奇异矩阵，则存在正交（酉）矩阵 Q 和实（复）非奇异上三角矩阵 R，使 A 有 QR 分解式（4.2.7）；且除去相差一个对角元素的绝对值（模）全等于 1 的对角矩阵因子外，分解式（4.2.7）是唯一的．

3.步骤

（1）写出矩阵的列向量；

（2）schmidt 正交化；

（3）得到 QR 分解结果:

Q=标准化的正交化结果所构成的矩阵，R= 对角标准化参数*计算参数。

4.2.2 举例展示求法

将矩阵 A $\begin{matrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 1 & 2 & 1 \end{matrix}$ 标准化的过程如下：

列向量 $\alpha_1=(1,2,1)^T,\ \alpha_2=(2,1,2)^T,\ \alpha_3=(2,2,1)^T$

正交得 $\beta_1=\alpha_1=(1,2,1)^T.$

$$\beta_2=\alpha_2-\beta_1\frac{(\alpha_2,\beta_1)}{(\beta_1,\beta_1)}=(1,-1,1)^T.$$

$$\beta_3=\alpha_3-\beta_1\frac{(\alpha_3,\beta_1)}{(\beta_1,\beta_1)}-\beta_2\frac{(\alpha_3,\beta_2)}{(\beta_2,\beta_2)}=(\tfrac{1}{2},0,-\tfrac{1}{2})^T.$$

构造矩阵，

$$Q=\text{标准化的}[\beta_1,\beta_2,\beta_3]=\begin{bmatrix}\frac{1}{\sqrt6}, & \frac{1}{\sqrt3} & \frac{1}{\sqrt2}\\ \frac{2}{\sqrt6} & -\frac{1}{\sqrt3} & 0\\ \frac{1}{\sqrt6} & \frac{1}{\sqrt3} & -\frac{1}{\sqrt2}\end{bmatrix}$$

$$R=\text{对角标准化参数}\times\text{计算参数}=\begin{bmatrix}\sqrt6 & 0 & 0\\ 0 & \sqrt3 & 0\\ 0 & 0 & \frac{1}{\sqrt2}\end{bmatrix}\begin{bmatrix}1 & 1 & \frac{7}{6}\\ 0 & 1 & \frac{1}{3}\\ 0 & 0 & 1\end{bmatrix}$$

$$A=QR.$$

4.3 矩阵的满秩分解

4.3.1 此处为矩阵的满秩分解的步骤推导

1.引入

（1）秩

设 A 是一组向量，定义 A 的极大无关组中向量的个数为 A 的秩。矩阵 A 的列秩是

A 的线性独立的纵列的极大数目，通常表示为 r(A)或 rank A。
满秩：m × n 矩阵的秩为 m 和 n 中的较小者。

（2）满秩分解

对于 m×n 的矩阵 A，假设其秩为 r，若存在秩同样为 r 两个矩阵：Cm×r（列满秩）和 Dr×n（行满秩），使得 A=CD，则称其为矩阵 A 的满秩分解。

（3）满秩分解不唯一，任何非零矩阵一定存在满秩分解。

2.步骤

（1）将 A 乘$L1^{(-1)}$，将第一列化为行阶梯形的最终形态，等到$A^{(1)}$；重复步骤，得到$A^{(1)}$…，$A^{(n)}$，$L1^{(-1)}$,...，$Ln^{(-1)}$A 化为行阶梯形矩阵$A^{(n)}$。

（2）A=L1…Ln$A^{(n)}$

4.3.2 举例展示求法

$$A=\begin{bmatrix}-1&0&1&2\\1&2&-1&1\\2&2&-2&-1\end{bmatrix}，求A=FG$$

1、将A逐步化为行阶梯形.

$$A \xrightarrow{r} A^{(1)} \xrightarrow{r} A^{(2)}$$

$$A^{(1)}=\begin{bmatrix}-1&0&1&2\\0&2&0&3\\0&2&0&3\end{bmatrix}=L_1^{-1}A.\quad A^{(2)}=\begin{bmatrix}-1&0&1&2\\0&2&0&3\\0&0&0&0\end{bmatrix}=L_2^{-1}A^{(1)}$$

$$L_1^{-1}=\begin{bmatrix}1&0&0\\1&1&0\\2&0&1\end{bmatrix}\quad L_2^{-1}=\begin{bmatrix}1&0&0\\0&1&0\\0&-1&1\end{bmatrix}$$

2.

$$A=L_1L_2A^{(2)}=\begin{bmatrix}1&0&0\\-1&1&0\\-2&1&1\end{bmatrix}\begin{bmatrix}-1&0&1&2\\0&2&0&3\\0&0&0&0\end{bmatrix}=\begin{bmatrix}1&0\\-1&1\\-2&1\end{bmatrix}\begin{bmatrix}-1&0&1&2\\0&2&0&3\end{bmatrix}$$

$$=FG.$$

4.4 矩阵的奇异值分解

4.4.1 矩阵的奇异值分解的步骤推导

1.引入

(1)奇异值的定义

定义 4.11 设 $\mathbf{A}\in \mathbf{C}_r^{m\times n}(r>0)$，$\mathbf{A}^H\mathbf{A}$ 的特征值为

$$\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_r>\lambda_{r+1}=\cdots=\lambda_n=0$$

则称 $\sigma_i=\sqrt{\lambda_i}\ (i=1,2,\cdots,n)$ 为 **A** 的**奇异值**；当 **A** 为零矩阵时，它的奇异值都是 0.

（2）矩阵奇异值分解的存在性

定理 4.16 设 $\boldsymbol{A} \in \mathbf{C}_r^{m\times n}(r>0)$，则存在 m 阶酉矩阵 $\boldsymbol{U}$ 和 n 阶酉矩阵 $\boldsymbol{V}$，使得

$$\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{V} = \begin{bmatrix} \boldsymbol{\Sigma} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{bmatrix} = D \qquad (4.4.4)$$

其中 $\boldsymbol{\Sigma} = \mathrm{diag}(\sigma_1, \sigma_2, \cdots, \sigma_r)$，而 $\sigma_i (i = 1, 2, \cdots, r)$ 为矩阵 $\boldsymbol{A}$ 的全部非零奇异值.

2.奇异值分解的定义

对于一个秩为 r 的矩阵A_{m*m} ,必存在 m*m 的正交矩阵U_{m*m}， n*n 的正交矩阵V_{n*n}，m*n 的矩阵 $\boldsymbol{\Sigma}_{m\times n}$ ，使得 $\boldsymbol{A}_{m\times n} = \boldsymbol{U}_{m\times m}\boldsymbol{\Sigma}_{m\times n}\boldsymbol{V}_{n\times n}^T = \boldsymbol{U}_{m\times m}\begin{pmatrix} \boldsymbol{D}_{r\times r} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix}_{m\times n}\boldsymbol{V}_{n\times n}^T$ ，

其中 $\boldsymbol{D}_{r\times r} = \begin{pmatrix} \sqrt{\lambda_1} & & & \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ & & & \sqrt{\lambda_r} \end{pmatrix}_{r\times r}$ ，$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_r > 0$ 为 $\boldsymbol{A}^T\boldsymbol{A}$ 的 r 个非零特征值（从大到小排列）。

3.步骤

（1）利用 A 的转置乘以 A 得到矩阵$A^T A$，对这个矩阵进行特征值分解，得到的 n 个特征向量张成的 n*n 矩阵就是 V 矩阵；

（2）令$U_1 = AV_1\Sigma^{-1}$，则 $A = U_1\Sigma V_1^H$；

(3)将 U1 扩充为酉矩阵$U = [U_1 \vdots U_2]$.最终得到 $A = UDV^H$

4.奇异值分解的推导

$$A = U\Sigma V^T \Rightarrow A^T = V\Sigma^T U^T \Rightarrow A^T A = V\Sigma^T U^T U\Sigma V^T = V\Sigma^2 V^T$$

4.4.2 举例展示求法

若 A= $\begin{matrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{matrix}$，求$\boldsymbol{A} = \boldsymbol{UDV}^T$.

1. 特征值分解，求V矩阵.

$A^TA=\begin{bmatrix}1&0&1\\0&1&1\\1&1&2\end{bmatrix}=B.\quad |\lambda I-B|=\lambda(\lambda-1)(\lambda-3)$

$\lambda_1=3$ 时, $3I-B=\begin{bmatrix}2&0&-1\\0&2&-1\\-1&-1&1\end{bmatrix}$, $\xi_1=\begin{bmatrix}1\\1\\2\end{bmatrix}$.

$\lambda_2=1$ 时, $1I-B=\begin{bmatrix}0&0&-1\\0&0&-1\\-1&-1&-1\end{bmatrix}$, $\xi_2=\begin{bmatrix}1\\-1\\0\end{bmatrix}$

$\lambda_3=0$ 时, $0I-B=\begin{bmatrix}-1&0&-1\\0&-1&-1\\-1&-1&-2\end{bmatrix}$, $\xi_3=\begin{bmatrix}1\\1\\-1\end{bmatrix}$.

$r_A=2,\quad \Sigma=\begin{bmatrix}\sqrt{3}&0\\0&1\end{bmatrix}$

$V=\begin{bmatrix}\frac{1}{\sqrt{6}}&\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{3}}\\\frac{1}{\sqrt{6}}&-\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{3}}\\\frac{2}{\sqrt{6}}&0&-\frac{1}{\sqrt{3}}\end{bmatrix}.\quad V_1=\begin{bmatrix}-\frac{1}{\sqrt{6}}&\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{6}}&-\frac{1}{\sqrt{2}}\\\frac{2}{\sqrt{6}}&0\end{bmatrix}.$

2.

$U_1=AV_1\Sigma^{-1}=\begin{bmatrix}\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{2}}\\0&0\end{bmatrix}.$

3. 将 U_1 扩充为酉矩阵 $U=[U_1 \vdots U_2]$.

取 $U_2=\begin{bmatrix}0\\0\\1\end{bmatrix}$. 则 $U=\begin{bmatrix}\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}&0\\\frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{2}}&0\\0&0&1\end{bmatrix}$.

$U^TAV=\begin{bmatrix}\sqrt{3}&0&0\\0&1&0\\0&0&0\end{bmatrix}=D,\quad A=UDV^T.$

4.4.3 利用奇异值分解求矩阵广义逆

1.引入

（1）广义逆矩阵定义

定义 6.4　设矩阵 $A\in C^{m\times n}$，若矩阵 $X\in C^{n\times m}$ 满足如下四个 Penrose 方程

$$AXA=A \tag{i}$$

$$XAX=X \tag{ii}$$

$$(AX)^H=AX \tag{iii}$$

$$(XA)^H=XA$$

则称 X 为 A 的 Moore-Penrose 逆，记为 A^+.

广义逆矩阵存在且唯一。

（2）

设 $A\in C_r^{m\times n}$ 的奇异值分解为 $A=U\begin{bmatrix}D_r & 0\\ 0 & 0\end{bmatrix}V=UDV$.

则有 $A^+=V^HD^+U^H$.

2.步骤

a)将矩阵进行奇异值分解

b)根据 $A^+=V^HD^+U^H$ 写出广义逆矩阵

3.举例

$$A=\begin{bmatrix}-1 & 0 & 1\\ 2 & 0 & -2\end{bmatrix}.$$

1. 求A的奇异值分解.

$$AA^H=\begin{bmatrix}2 & -4\\ -4 & 8\end{bmatrix}=\begin{bmatrix}\frac{1}{\sqrt5} & \frac{2}{\sqrt5}\\ -\frac{2}{\sqrt5} & \frac{1}{\sqrt5}\end{bmatrix}\begin{bmatrix}10 & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt5} & -\frac{2}{\sqrt5}\\ \frac{2}{\sqrt5} & \frac{1}{\sqrt5}\end{bmatrix}$$

令 $V_1=\frac{1}{\sqrt{10}}\begin{bmatrix}\frac{1}{\sqrt5} & -\frac{2}{\sqrt5}\end{bmatrix}\begin{bmatrix}-1 & 0 & 1\\ 2 & 0 & -2\end{bmatrix}=\begin{bmatrix}-\frac{1}{\sqrt2} & 0 & \frac{1}{\sqrt2}\end{bmatrix}$

$$A=\begin{bmatrix}\frac{1}{\sqrt5}\\ -\frac{2}{\sqrt5}\end{bmatrix}\begin{bmatrix}\sqrt{10}\end{bmatrix}\begin{bmatrix}-\frac{1}{\sqrt2} & 0 & \frac{1}{\sqrt2}\end{bmatrix}$$

2. 写出广义逆矩阵

$$A^+=\begin{bmatrix}-\frac{1}{\sqrt2}\\ 0\\ \frac{1}{\sqrt2}\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{10}}\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt5} & -\frac{2}{\sqrt5}\end{bmatrix}=\frac{1}{10}\begin{bmatrix}-1 & 2\\ 0 & 0\\ 1 & -2\end{bmatrix}$$

第五章 总结

本论文分别研究了矩阵函数的求法和矩阵分解方法。矩阵函数的求法和矩阵分解方法是矩阵论的精华知识之一，涉及大量矩阵论知识，需要对之前学过的知识进行融会贯通。经过本次研究学习，我加深了对线性变换、矩阵函数的理解，能够熟练地选用合适的方法进行矩阵函数求解和矩阵分解。同时，通过查阅相关资料，我更加意识到矩阵论不仅是一门停留在书本上的学科，它是更新着的、与计算机紧密融合的学科。通过对矩阵的收敛性、矩阵级数、矩阵函数、矩阵微分、矩阵积分、矩阵四种分解等系统性学习研究，让我明白了矩阵理论在实际生活中的巨大作用——矩阵论将大大减少工程运算量及提高计算速度、精度。有了矩阵理论作指导，现实生活中很多不能解决或者很难解决的数学问题等都能够得到很好的解决。比如，提高计算机的计算速度、优化数字信号处理算法等。今后我也会继续保持在矩阵领域的探索与学习，将其化为自己改进算法、拓展思路的有利工具。

参考文献

[1] 张凯院，徐仲等，矩阵论，西北工业大学出版社，2017 年 8 月.
[2] J. Shen, T. Tang, Spectral and High-orderMethods with Applications, Science Press, 2006.