

**期 末 课 程 论 文**



**题目: 矩阵函数的求法和矩阵分解方法研究**

**姓 名 陈子容**

**学 院 计算机学院**

**专 业 计算机类**

**班 级 2023211320**

**学 号 2023210710**

**指导教师 李昊辰**

# **2024年 12 月**

**矩阵函数的求法和矩阵分解方法研究**

**摘 要**

本文深入探讨了矩阵函数求解和矩阵分解方法，为矩阵论及其应用领域提供了重要的理论基础和实用工具。首先，文章详细介绍了线性空间、线性变换、欧氏空间等基本概念，并探讨了Jordan标准形及其求解方法，为后续矩阵函数的研究奠定了基础。接着，文章深入研究了向量范数、矩阵范数以及相关概念，并将其应用于K-means聚类算法中。关于矩阵函数，文章不仅阐述了其定义、性质，还详细介绍了利用待定系数法、数项级数求和法、对角型法和Jordan标准形法求解矩阵函数的方法，并通过实例验证了这些方法的有效性。此外，本文还系统地研究了矩阵的LU分解、QR分解、满秩分解以及奇异值分解等重要的矩阵分解方法，并通过实例展示了其在求解线性方程组、图像压缩等领域的应用。最后，文章深入探讨了矩阵广义逆的概念及其利用矩阵分解求解的方法，进一步完善了矩阵理论。总而言之，本文系统性地介绍了矩阵函数与矩阵分解的相关理论，并通过实例和代码验证了理论的有效性，为相关领域的研究提供了有价值的参考。

**关键词** 矩阵论 矩阵函数 矩阵分解

**Research on the Methods of Matrix Function Calculation and Matrix Decomposition**

**ABSTRACT**

The text delves deeply into the methods of solving matrix functions and matrix decomposition, providing an important theoretical foundation and practical tools for matrix theory and its application fields. Firstly, the article introduces fundamental concepts such as linear spaces, linear transformations, and Euclidean spaces in detail, and explores the Jordan canonical form and its solving methods, laying the groundwork for subsequent research on matrix functions. Next, the article conducts an in-depth study of vector norms, matrix norms, and related concepts, applying them to the K-means clustering algorithm. Regarding matrix functions, the article not only elucidates their definitions and properties but also provides detailed methods for solving matrix functions using the method of undetermined coefficients, series summation, diagonalization, and the Jordan canonical form, validating the effectiveness of these methods through examples. Furthermore, this paper systematically investigates important matrix decomposition methods such as LU decomposition, QR decomposition, full-rank decomposition, and singular value decomposition, demonstrating their applications in solving linear equations and image compression through examples. Finally, the article thoroughly discusses the concept of the generalized inverse of a matrix and the methods for solving it using matrix decomposition, further enhancing matrix theory. In summary, this paper systematically introduces the relevant theories of matrix functions and matrix decomposition, validating the effectiveness of the theories through examples and code, providing valuable references for research in related fields.

**KEY WORDS** Matrix theory Matrix functions Matrix decomposition

**目 录**

[第一章 引言 1](#_Toc21584)

[1.1 背景介绍 1](#_Toc4209)

[1.1.1 矩阵理论与方法介绍 1](#_Toc29763)

[1.1.2 函数矩阵和矩阵函数介绍 1](#_Toc28573)

[1.1.3 线性代数方程组求解介绍 1](#_Toc31437)

[1.2 问题介绍 1](#_Toc19570)

[1.2.1 矩阵函数的求法问题介绍 1](#_Toc32017)

[1.2.2 矩阵分解的方法问题介绍 1](#_Toc15611)

[1.3 上述问题国内外研究成果介绍 2](#_Toc155)

[1.3.1 矩阵函数的求法研究现状 2](#_Toc403)

[1.3.2 矩阵分解方法研究现状 2](#_Toc18340)

[1.4 本论文工作简述 2](#_Toc25796)

[1.4.1 本论文对上述问题研究简述 2](#_Toc26019)

[1.4.2 本论文创新点或特点简述 2](#_Toc30541)

[1.4.3 本论文撰写结构简述 3](#_Toc3540)

[第二章 预备知识 4](#_Toc1147)

[2.1 欧氏空间与线性变换 4](#_Toc12651)

[2.1.1 欧氏空间与线性变换介绍 4](#_Toc5821)

[2.1.2 Jordan标准形的求解 6](#_Toc27379)

[2.1.3 欧氏空间中线性变换的求法 8](#_Toc25048)

[2.2 向量范数与矩阵范数 10](#_Toc26716)

[2.2.1 向量范数介绍 10](#_Toc28667)

[2.2.2 矩阵范数介绍 12](#_Toc16525)

[2.2.3 矩阵可逆性条件、条件数和谱半径介绍 13](#_Toc25663)

[2.3 矩阵函数介绍 13](#_Toc13656)

[2.3.1 矩阵序列介绍 13](#_Toc11164)

[2.3.2 矩阵级数介绍 14](#_Toc30477)

[2.3.3 矩阵函数介绍 14](#_Toc14229)

[2.3.4 函数矩阵对矩阵的导数 15](#_Toc5732)

[第三章 矩阵函数的求法研究 19](#_Toc9280)

[3.1 待定系数法 19](#_Toc32041)

[3.1.1 待定系数法求矩阵函数的步骤推导 19](#_Toc19826)

[3.1.2 举例展示求法 19](#_Toc23713)

[3.2 数项级数求和法 19](#_Toc7163)

[3.2.1 数项级数求和法求矩阵函数的步骤推导 19](#_Toc23599)

[3.2.2 举例展示求法 20](#_Toc21670)

[3.3 对角形法 21](#_Toc5626)

[3.3.1 对角形法求矩阵函数的步骤推导 21](#_Toc15006)

[3.3.2 举例展示求法 21](#_Toc27499)

[3.4 Jordan标准形法 22](#_Toc7581)

[3.4.1 Jordan标准形法求矩阵函数的步骤推导 22](#_Toc10161)

[3.4.2 举例展示求法 24](#_Toc8916)

[第四章 矩阵分解方法研究 24](#_Toc24400)

[4.1 矩阵的LU分解 25](#_Toc5081)

[4.1.1 矩阵LU分解的步骤推导 25](#_Toc27443)

[4.1.2 举例展示求法 28](#_Toc20210)

[4.2 矩阵的QR分解 29](#_Toc4552)

[4.2.1 矩阵QR分解的步骤推导 30](#_Toc4095)

[4.2.2 举例展示求法 31](#_Toc11642)

[4.3 矩阵的满秩分解 33](#_Toc1417)

[4.3.1 矩阵满秩分解的步骤推导 33](#_Toc12961)

[4.3.2 举例展示求法 34](#_Toc2186)

[4.4 矩阵的奇异值分解 35](#_Toc14317)

[4.4.1 矩阵奇异值分解的步骤推导 35](#_Toc2369)

[4.4.2 举例展示求法 36](#_Toc8990)

[4.5 利用矩阵分解求矩阵广义逆 36](#_Toc16467)

[4.5.1 矩阵广义逆介绍 40](#_Toc28413)

[4.5.2 利用矩阵满秩分解求矩阵广义逆 41](#_Toc11408)

[4.5.3 利用矩阵奇异值分解求矩阵广义逆 41](#_Toc32686)

[4.5.4 举例展示求法 42](#_Toc4410)

[第五章 总结 43](#_Toc12365)

[参考文献 47](#_Toc19412)

**第一章 引言**

## **1.1 背景介绍**

### **1.1.1 矩阵理论与方法介绍**

矩阵理论是研究矩阵及其运算、性质和应用的数学分支，涉及矩阵的加法、乘法、转置、逆矩阵、行列式等基本操作。它广泛应用于解决线性方程组、优化问题、特征值与特征向量分析、数据处理等领域。矩阵分解（如LU分解、QR分解、SVD分解）是其重要方法，常用于提高计算效率和求解复杂问题。矩阵理论在工程、物理学、计算机科学、经济学等多个学科中发挥着关键作用。

### **1.1.2 函数矩阵和矩阵函数介绍**

函数矩阵和矩阵函数是矩阵理论中的重要概念。函数矩阵指的是矩阵中的元素是某个函数的形式，比如矩阵的每个元素都可以是一个标量函数。矩阵函数则是将矩阵作为自变量，定义一个类似于标量函数的操作。例如，矩阵的指数、对数、平方根等可以被看作是矩阵函数。这些概念在解决矩阵方程、控制理论及数值分析中有广泛应用。

### **1.1.3 线性代数方程组求解介绍**

线性代数方程组求解是矩阵理论中的基础问题，通常形式为，其中是系数矩阵，是未知向量，是常数向量。常见的求解方法包括高斯消元法、矩阵的逆、克拉默法则、以及LU分解等。这些方法在工程、物理、经济学等领域具有广泛应用。

## **1.2 问题介绍**

### **1.2.1 矩阵函数的求法问题介绍**

矩阵函数的求解涉及将矩阵作为自变量，定义类似于标量函数的操作。常见的矩阵函数包括矩阵的指数、对数、平方根、幂函数等。求解矩阵函数通常依赖于矩阵的谱分解、特征值分解或Jordan标准型等方法。具体方法包括通过对角化矩阵、使用泰勒级数展开、或利用数值算法进行近似计算。这些方法在控制理论、量子力学、系统工程等领域有重要应用。

### **1.2.2 矩阵分解的方法问题介绍**

矩阵分解方法是将一个矩阵分解为多个简单矩阵的乘积，以便简化计算和分析。常见的矩阵分解方法包括：

LU分解：将矩阵分解为下三角矩阵和上三角矩阵的乘积，用于求解线性方程组。

QR分解：将矩阵分解为正交矩阵和上三角矩阵的乘积，常用于最小二乘问题。

满秩分解：将矩阵分解为两个满秩矩阵的乘积。

奇异值分解：将任意矩阵分解为三个矩阵的乘积，广泛用于数据降维、信号处理等领域。

这些分解方法帮助解决矩阵的逆、方程求解、数据分析等问题，具有重要的理论和实际意义。

##### **1.3 上述问题国内外研究成果介绍**

# **1.3.1 矩阵函数的求法研究现状**

矩阵函数的求法研究现状主要集中在提高计算效率和精度方面，特别是在矩阵指数、对数、平方根和幂等函数的计算上。常用的求解方法包括泰勒级数展开、Padé逼近、谱分解法、Schur分解法和迭代法等。此外，研究还着眼于稀疏矩阵、并行计算和GPU加速等技术，以适应大规模数据处理的需求。随着数值算法和计算资源的进步，矩阵函数的计算在控制理论、量子力学、机器学习等领域的应用得到了广泛关注和深入发展。

# **1.3.2 矩阵分解方法研究现状**

矩阵分解方法的研究现状主要集中在提高计算效率、精度和适应性方面。常见的矩阵分解方法包括奇异值分解(SVD)、特征值分解、LU分解、QR分解、Cholesky分解以及低秩分解等。近年来，针对大规模矩阵的计算，研究者还提出了基于稀疏矩阵、分布式计算和并行计算的优化算法。此外，矩阵分解在数据压缩、机器学习、信号处理和图像处理等领域得到了广泛应用，尤其是在处理大规模数据和高维问题时，矩阵分解方法的高效性和稳定性成为研究的重点。

##### **1.4 本论文工作简述**

# **1.4.1 本论文对上述问题研究简述**

本论文旨在研究矩阵函数的求法与矩阵分解方法的有效应用。对于矩阵函数的求法，本文重点分析了一般矩阵函数的数值计算方法，包括待定系数法、数项级数求和法、对角形法以及Jordan标准形法等，并探讨了这些方法在不同矩阵类型上的适用性与计算效率。

在矩阵分解方法方面，本文详细讨论了传统矩阵分解方法（如LU分解、QR分解、奇异值分解等）的理论背景和计算实现，并探索了它们在解决大规模线性方程组、数据压缩等领域中的应用。

# **1.4.2 本论文创新点或特点简述**

在本文中，本文提出了以下几个创新点：

1. 多项式线性空间与矩阵空间在数值计算中的应用：

本文首先介绍了多项式线性空间和矩阵空间在数值分析和计算机科学中的广泛应用。同时，矩阵空间在图像、文本和声音等数据处理中的应用也被详细阐述，展示了线性空间概念在实际应用中的广泛性和重要性。

1. 欧氏空间中线性变换的求法：

本文提出了一种通过求解欧氏空间中线性变换的标准正交基、对称变换和对角矩阵表示的方法。这种方法不仅简化了线性变换的计算，还为后续的矩阵分解和函数求解提供了理论基础。

1. 矩阵函数求解的多种方法：

本文介绍了待定系数法、数项级数求和法、对角型法和Jordan标准形法来求解矩阵函数。每种方法都有其独特的优势和适用范围，特别是对于与对角矩阵相似的矩阵，对角型法和Jordan标准形法显著提高了计算效率。

1. 矩阵分解方法的多种应用：

本文详细介绍了LU分解、QR分解、满秩分解和奇异值分解，并展示了这些分解方法在线性方程组求解、图像压缩、3D变换分解等实际应用中的重要性。特别是，奇异值分解（SVD）在图像压缩、医学成像和视频压缩等领域的应用，展示了矩阵分解在实际问题中的广泛应用和重要性。

1. 矩阵广义逆的概念与应用：

本文引入了矩阵广义逆的概念，并展示了其在线性方程组求解中的重要性。本文提出了利用矩阵满秩分解和奇异值分解来求解矩阵广义逆的方法，并验证了这些方法在实际应用中的有效性和高效性。

1. 反向传播算法中函数矩阵的导数应用：

本文展示了函数矩阵的导数在深度学习中反向传播算法中的应用。通过将复杂的代数式化为更简洁的形式，本文能更容易地观察到问题的关键，从而给出准确且高效的解决方案。

这些创新点不仅丰富了矩阵分析的理论内容，还展示了这些理论在实际应用中的广泛和深远的影响。

# **1.4.3 本论文撰写结构简述**

本论文分为五章，具体结构如下：

第一章 引言（本章）：介绍了矩阵理论和矩阵函数的基本概念，阐述了矩阵函数求法和矩阵分解方法的研究背景和现状，以及论文的研究目标和意义。

第二章 预备知识：本章回顾了欧氏空间与线性变换的基础知识，并对矩阵范数、条件数等相关理论进行了详细介绍，为后续矩阵函数与矩阵分解方法的讨论提供理论支持。

第三章 矩阵函数的求法研究：本章系统地介绍了几种常用的矩阵函数求解方法，包括待定系数法展开法、数项级数求和法、对角形法和Jordan标准形法等，并通过具体例子展示了这些方法的实现。

第四章 矩阵分解方法研究：本章深入探讨了矩阵LU分解、QR分解、奇异值分解等经典矩阵分解方法，详细介绍了它们的数学原理和数值实现，特别是它们在不同领域的应用。

第五章 总结：总结了本论文的主要研究成果，回顾了矩阵函数求法和矩阵分解方法的优化策略，提出了未来研究的可能方向和改进空间。

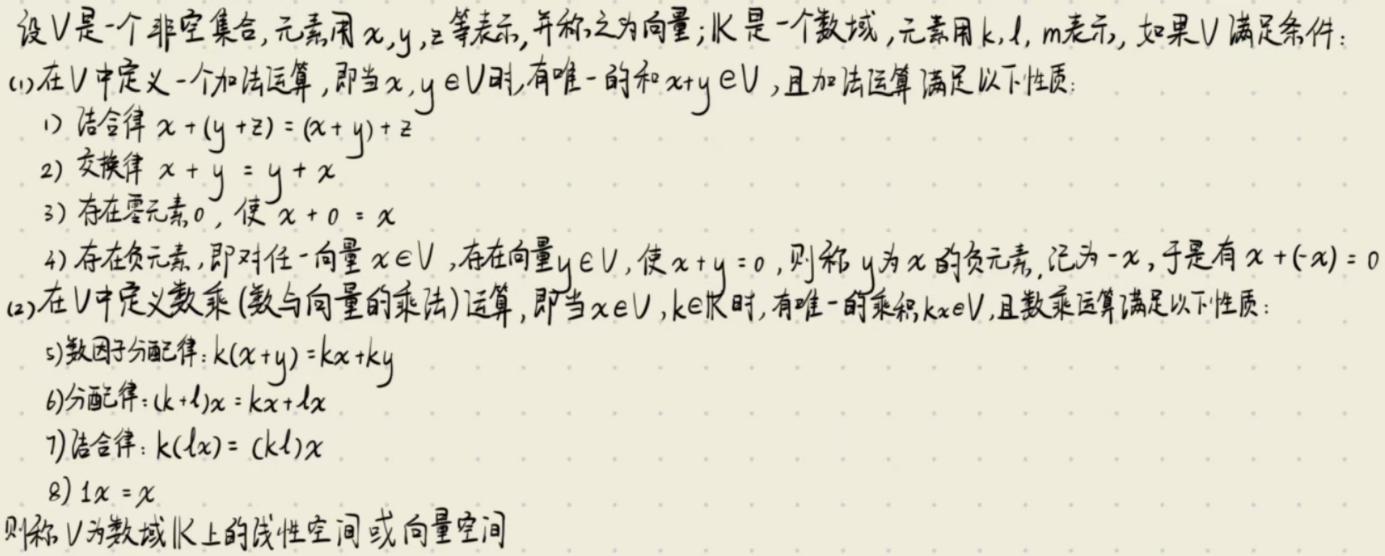
通过这种结构安排，本论文不仅全面系统地讨论了矩阵函数求法和矩阵分解方法的理论与实践问题，同时也为进一步的研究提供了可操作的思路和方法。

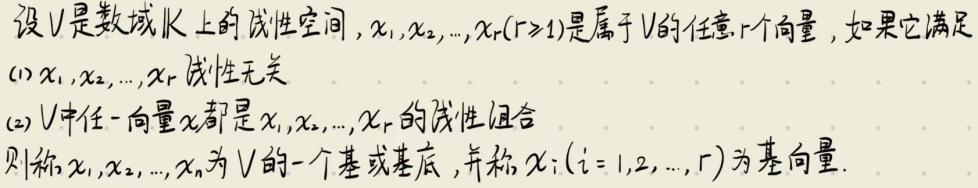
# **第二章 预备知识**

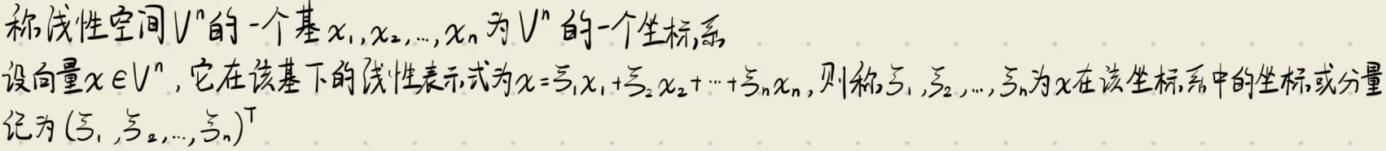
##### **2.1 欧氏空间与线性变换**

# **2.1.1 欧氏空间与线性变换介绍**

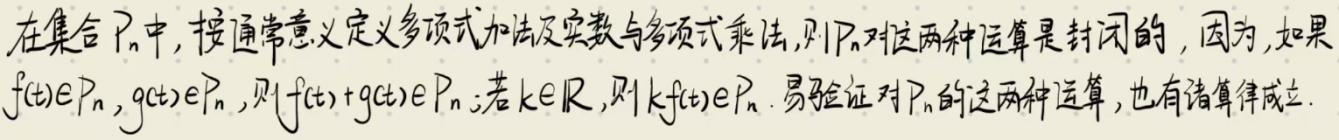
1. 线性空间相关概念



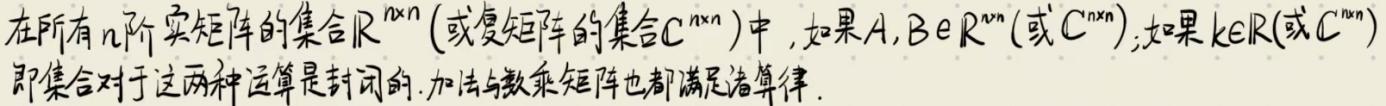




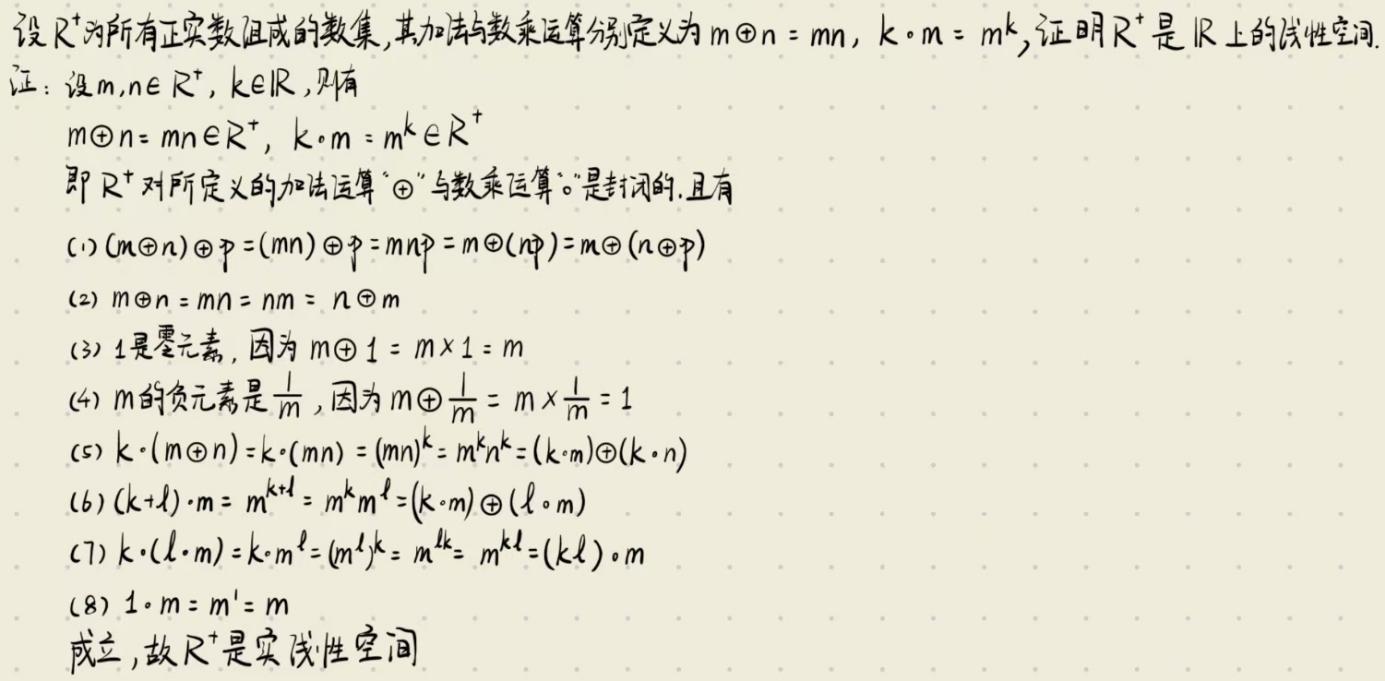
1. 线性空间举例



上述线性空间常称为多项式线性空间。多项式线性空间在数值分析中常用于函数近似。具体来说，许多函数可以通过多项式进行近似，特别是在数值计算中。泰勒级数就是一种通过多项式来近似某些函数的方法。在这种情境下，多项式空间成为描述和近似连续函数的重要工具。

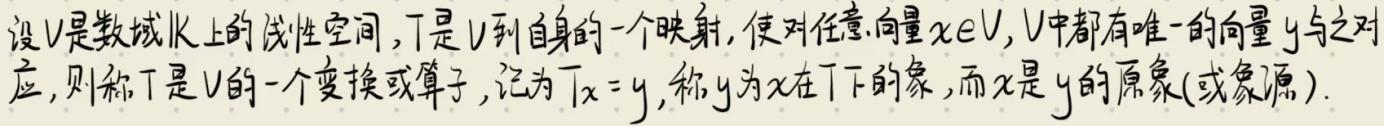


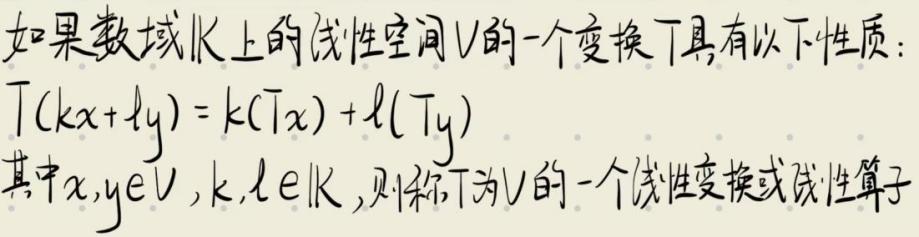
上述线性空间常称为矩阵空间。在计算机图形学，机器学习和数据科学中，图像、文本和声音等数据通常以矩阵的形式存储和处理。比如，一张图像可以表示为一个二维矩阵，其中每个元素表示像素值。在这种情境下，矩阵空间成为表示和操作图像的关键工具。

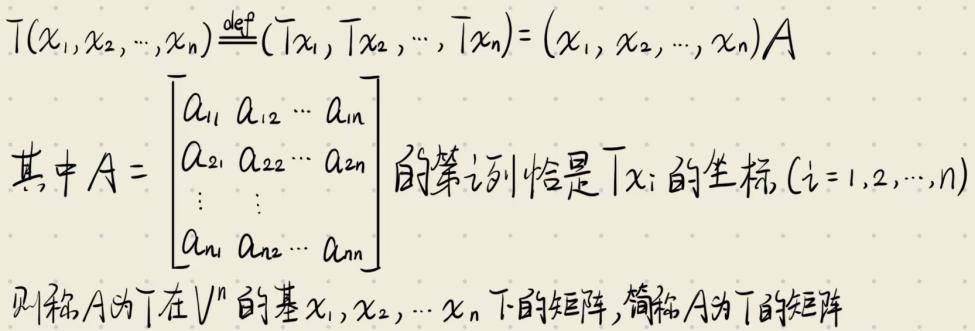


上例表明，线性空间的概念是十分广泛的，线性空间中加法和数乘的定义也并非局限于通常的定义。

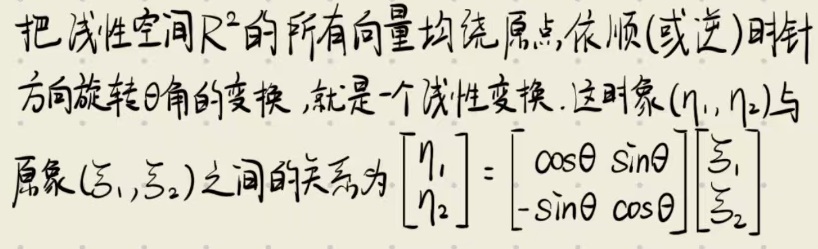
1. 线性变换相关概念



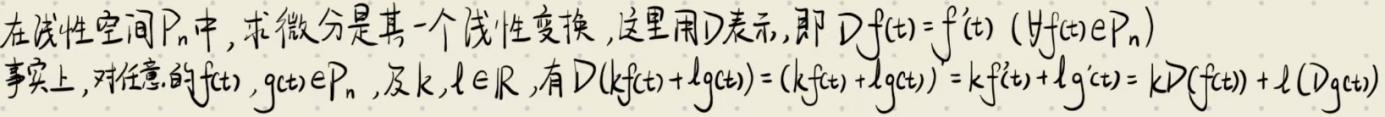


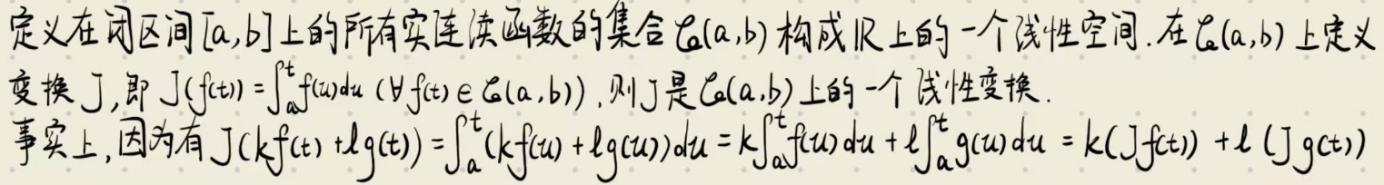


1. 线性变换举例



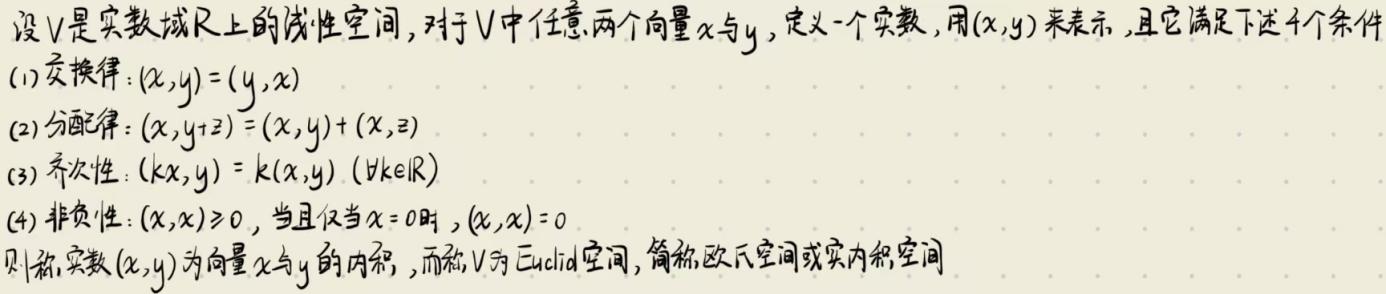
上述变换是计算机图形学中旋转变换实现的理论基础





上述两变换表明多项式的求导与积分是线性运算，也是计算机计算导数与积分的理论基础

1. 欧氏空间的定义

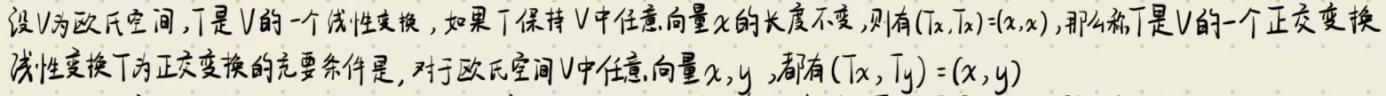


在复线性空间中，也有类似的酉空间的定义，即复内积空间

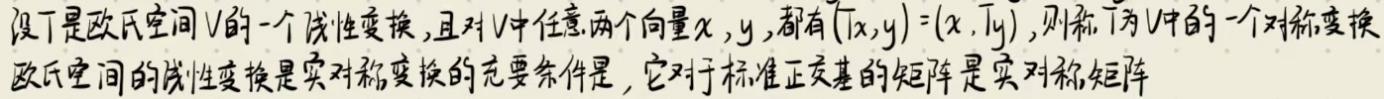
如果对于欧氏空间中的两个向量与，有，则称与正交

2.1.1.5.2

1. 欧氏空间中的线性变换



在酉空间中，也有类似的酉变换的定义，其在标准正交基下的矩阵称为酉矩阵



在酉空间中，也有类似的Hermite变换的定义，其在标准正交基下的矩阵称为Hermite矩阵

# **2.1.2 Jordan标准形的求解**

1. Jordan标准形的概念

若的特征多项式可分解因式为

则相似于以下形式的准对角矩阵

其中



1. 多项式矩阵理论介绍

多项式矩阵形式如下

其中为数域上的纯量的多项式

多项式矩阵的标准形，是指使用矩阵的初等变换将化为多项式矩阵，有

其中

且是首多项式（前面几个可能是），通常称为的不变因子。把的每个次数大于零的不变因子分解为不可约因式的乘积，这样的不可约因式（连同它们的幂指数）称为的一个初等因子，初等因子的全体称为的初等因子组

1. Jordan标准形的求法

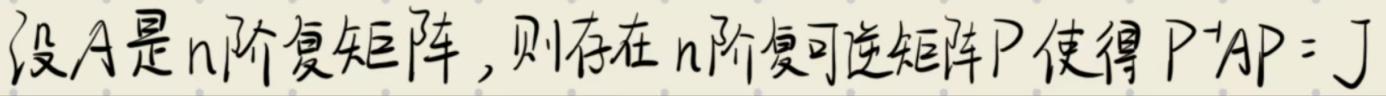
在复数域上，求阶矩阵的Jordan标准形的步骤如下：

第一步：求特征矩阵的初等因子组，设为

第二步：写出每个初等因子对应的Jordan块

第三步：写出以这些Jordan块构成的Jordan标准形

1. Jordan标准形的性质



上述性质表明，任意方阵都存在Jordan标准形，且任意方阵都与Jordan标准形相似，这使得任何处理方阵的问题都可以转化为处理其Jordan标准形的问题。而方阵的Jordan标准形又具有较为简单的分块对角的形式，这使得处理方阵的各种问题都可以由此变得更为简便。这一点在本文后续会得到充分印证。

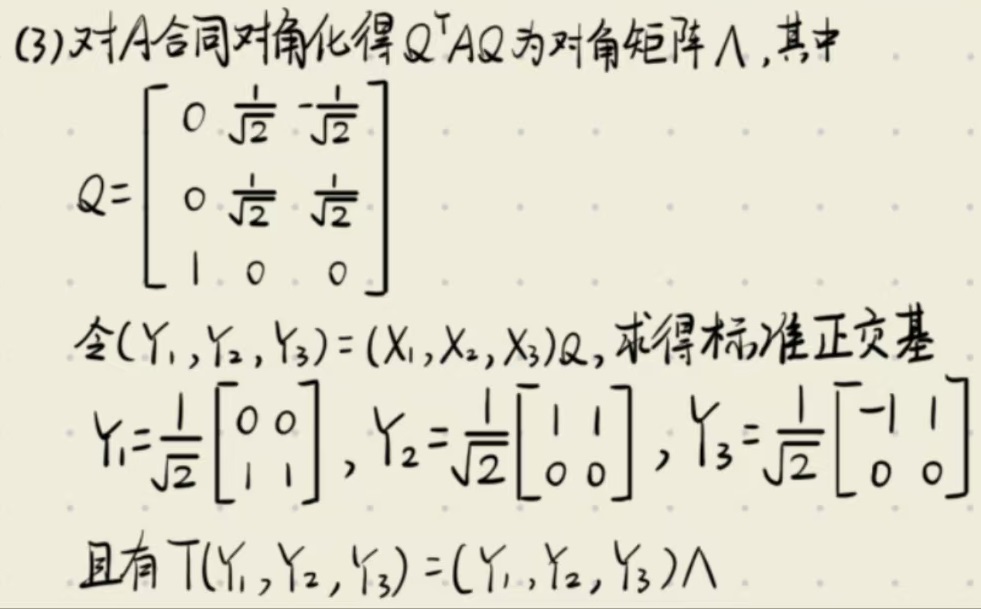
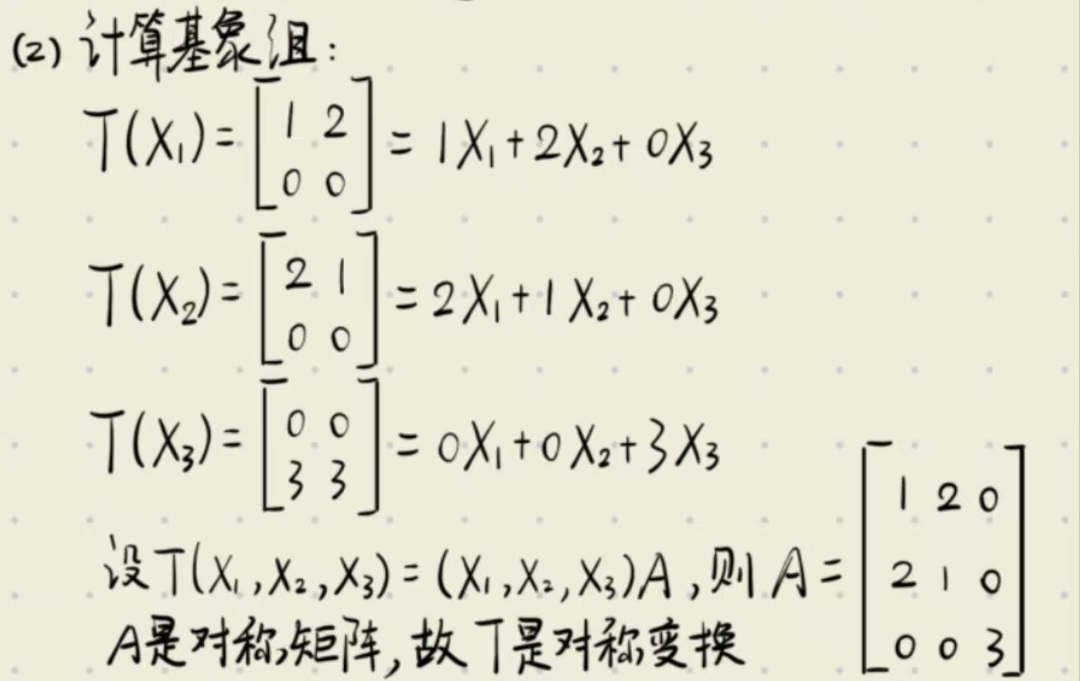
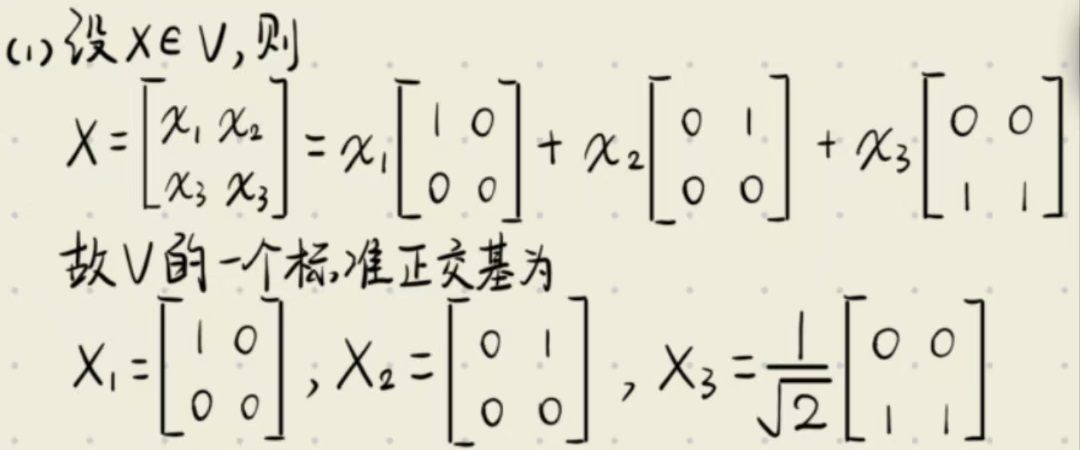
# **2.1.3 欧氏空间中线性变换的求法**

现以例题的形式介绍欧氏空间中线性变换的求法

在欧氏空间中，矩阵和的内积定义为，子空间

中的线性变换为

1. 求的一个标准正交基
2. 验证是中的对称变换
3. 求的一个标准正交基，使在该基下的矩阵为对角矩阵

****

现对上述解题过程一般化。对于欧氏空间中的任意对称变换，其在标准正交基下的矩阵为对称矩阵，故存在正交矩阵，使得为对角矩阵。以为基变换矩阵对原基做基变换得到新基并单位化，则新基也为标准正交基。故对于欧氏空间中的任意对称变换，总存在某个标准正交基，使其在该基下的矩阵为对角矩阵。

类似的，对于酉空间中的任意Hermite变换，总存在某个标准正交基，使其在该基下的矩阵为对角矩阵。

##### **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



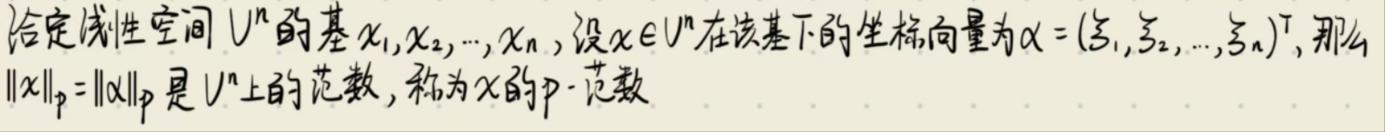
1. **-**范数介绍

对于不小于的任意实数及，可以证明实值函数

满足向量范数定义的三个条件，称为向量的**-**范数，记为，即

上式中令有

称为向量的**-**范数，记为



1. 向量范数等价

2.2.1.3

2.2.1.4

1. 向量范数的应用举例

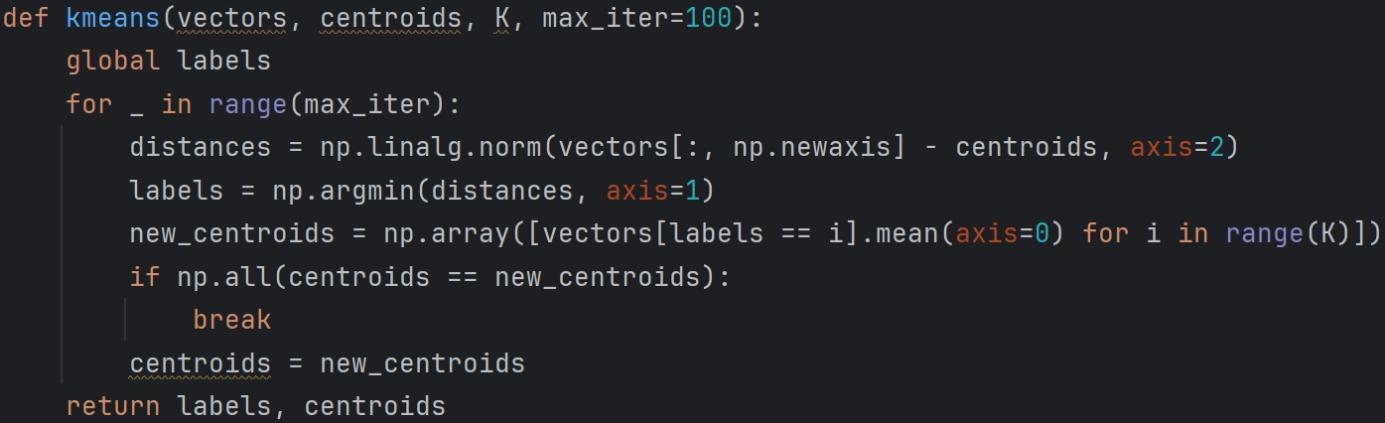
向量范数广泛应用与许多领域，这里仅举一例——K-means聚类算法

K-means聚类算法是一种常见的无监督学习算法，用于将数据集分成类，每个类包含具有相似特征的数据点。算法的核心思想是通过迭代优化类的划分，使得类内的数据点尽可能相似，类间的数据点尽可能不同。

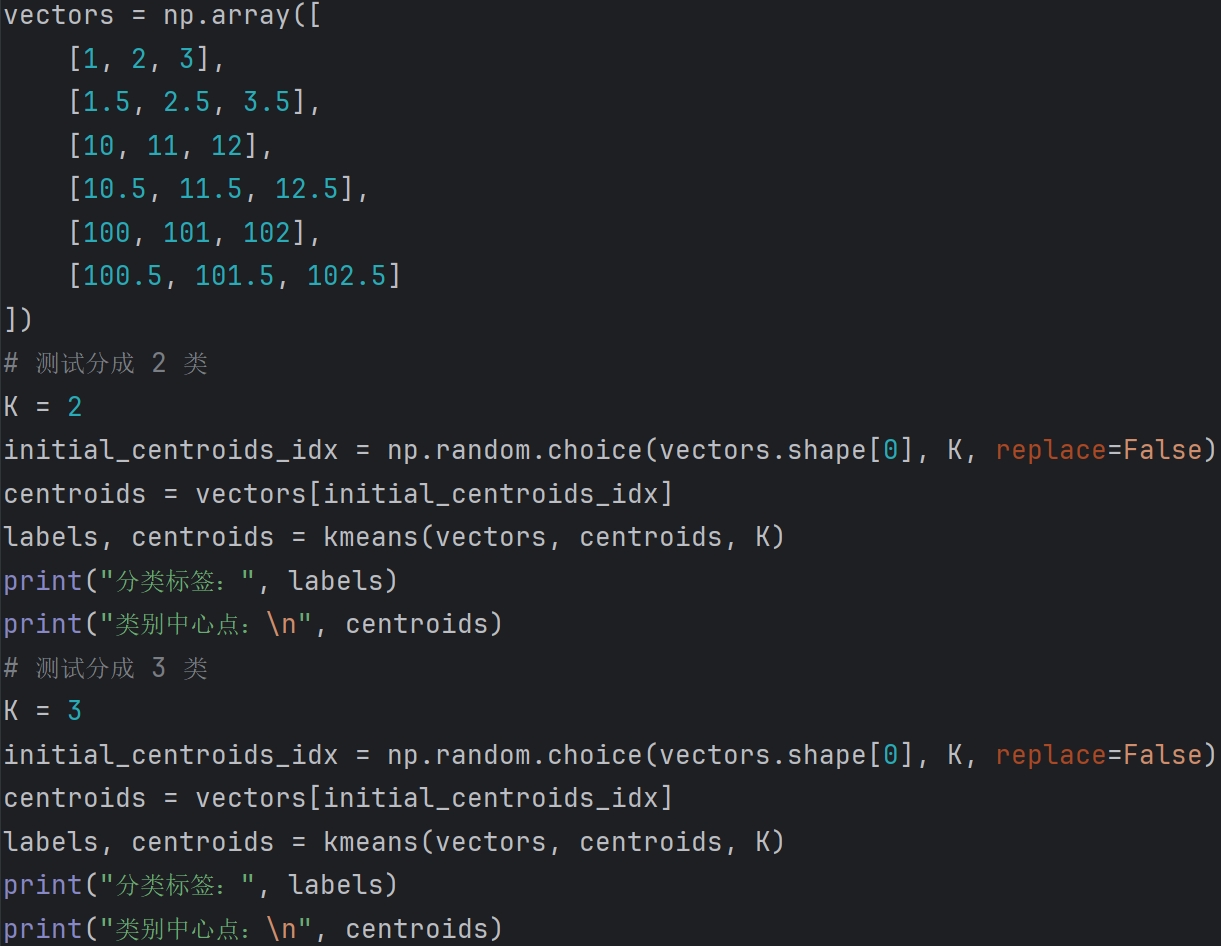
K-means算法的主要步骤如下：

* 1. 初始化类中心：随机选择个数据点作为初始的类中心。
  2. 分配数据点：对于数据集中的每个数据点，计算它与所有类中心的距离，通常使用向量范数中的**-**范数来度量距离，将每个数据点分配给距离最近的类中心。
  3. 更新类中心：根据分配的类中所有数据点的位置，计算每个类的质心，并将质心作为新的类中心。
  4. 迭代优化：重复步骤和，直到类中心不再发生显著变化，或者达到指定的迭代次数。

以下给出基于**-**范数的K-means聚类算法的代码实现



测试部分代码如下



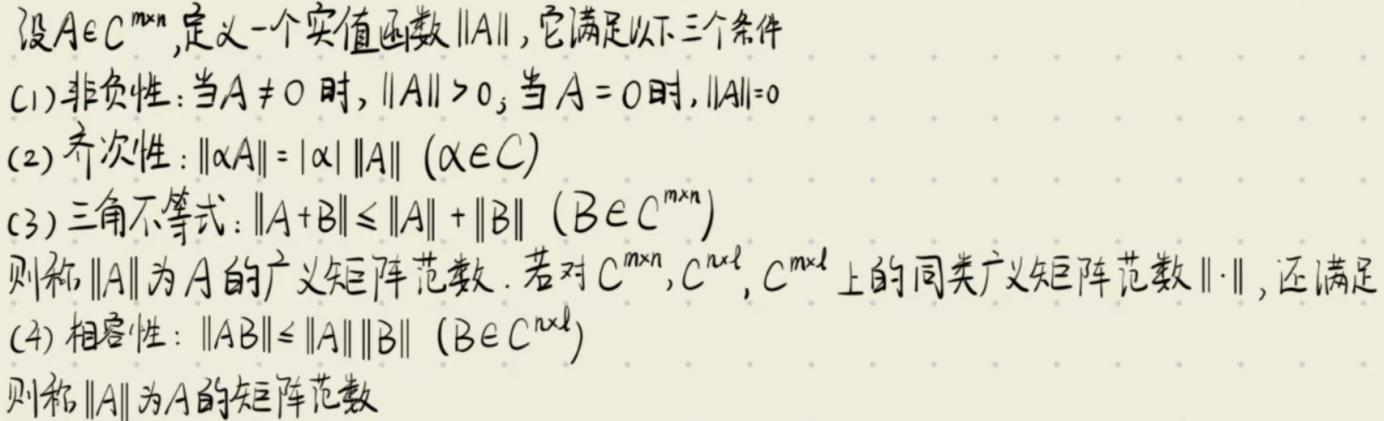
输出结果如下



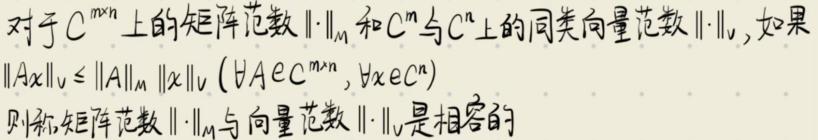
在K-means算法中，向量范数主要用于度量数据点与类中心之间的距离，在K-means算法中起到了核心作用，帮助算法完成数据点的分配和类中心的更新，使得聚类结果尽可能地符合预期目标。

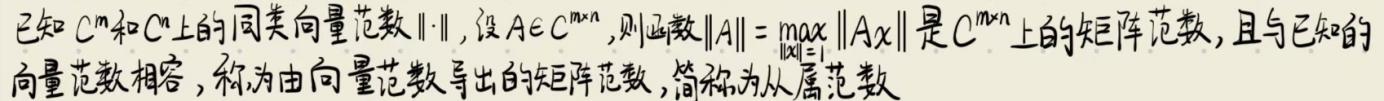
# **2.2.2 矩阵范数介绍**

1. 矩阵范数的定义

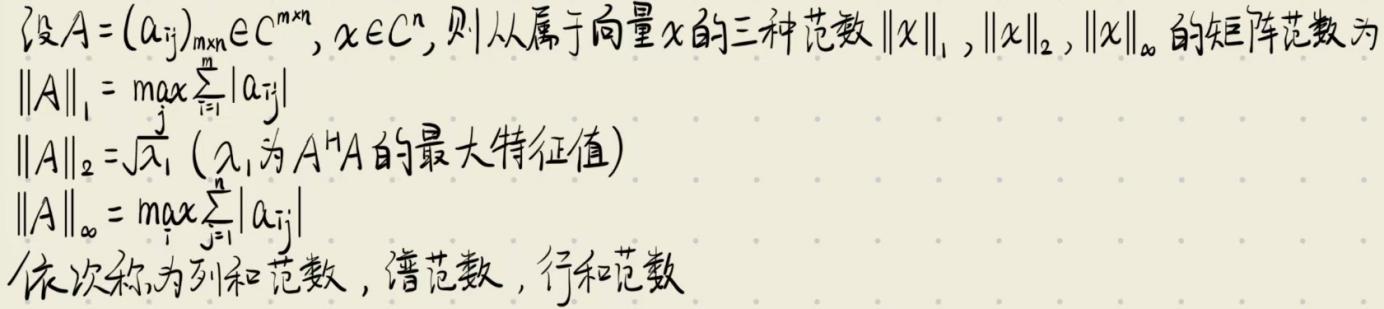


1. 矩阵范数与向量范数的关系





1. 常用矩阵范数介绍



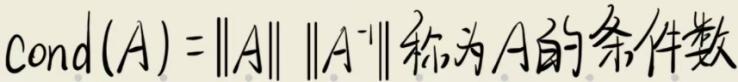
# **2.2.3 矩阵可逆性条件、条件数和谱半径介绍**

1. 矩阵可逆性条件介绍

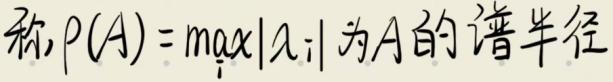
设，且对于上的某种矩阵范数，有，则矩阵可逆，且有

1. 条件数介绍

设的元素带有误差，则准确矩阵应为，其中。若为可逆矩阵，则用近似的相对误差满足

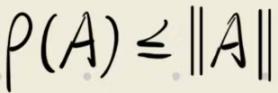
其中，，一般说来，条件数愈大，与的相对误差就愈大

1. 谱半径介绍

设的个特征值为，

的谱半径具有以下性质

* 1. 对上任何一种矩阵范数，都有

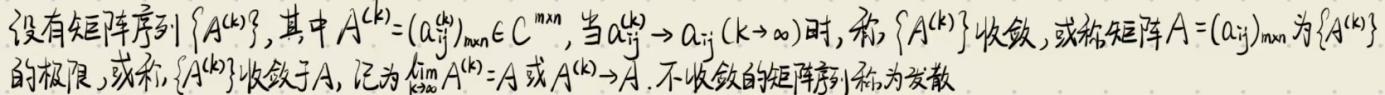


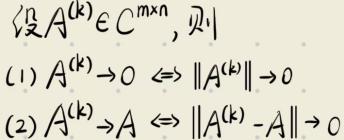
* 1. 

##### **2.3 矩阵函数介绍**

# **2.3.1 矩阵序列介绍**

1. 矩阵序列及收敛性





2.3.1.2

1. 收敛矩阵

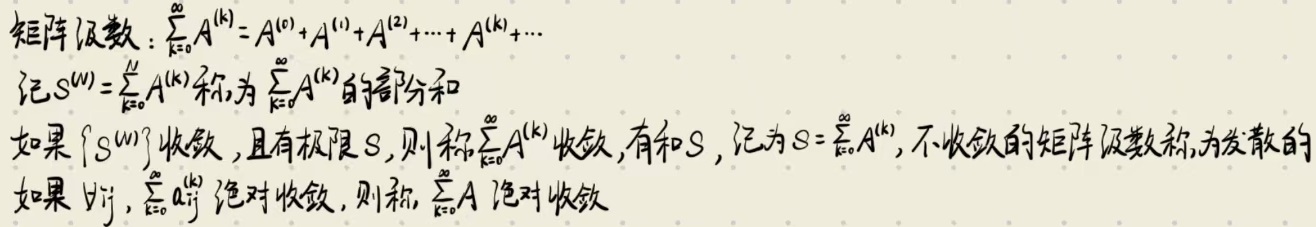
2.3.1.3

6c3c514aedf6867811d1719b8fdeb6c

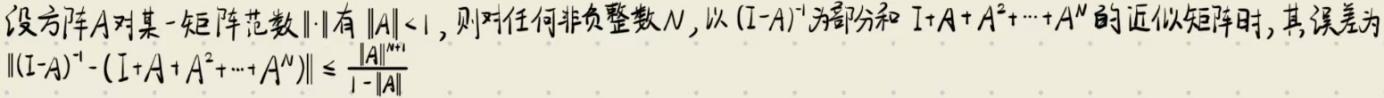
2.3.1.6

# **2.3.2 矩阵级数介绍**

1. 矩阵级数相关概念



1. 矩阵和估计相关定理



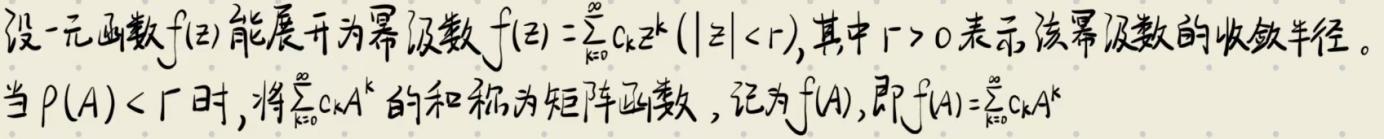
1. 从矩阵级数到矩阵函数

2.3.2.3

上述定理是定义矩阵函数的理论基础

# **2.3.3 矩阵函数介绍**

1. 矩阵函数的定义



1. 常用矩阵函数

称式

为矩阵指数函数

称式

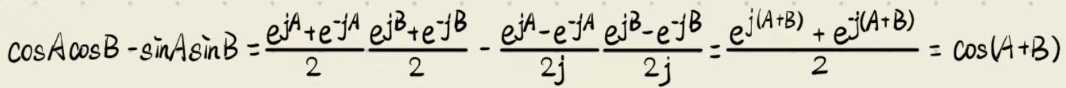
为矩阵三角函数

1. 常用矩阵函数的性质

2.3.3.2

由此可证得以下性质成立

这里只证第一个等式，其余等式证明方法类似

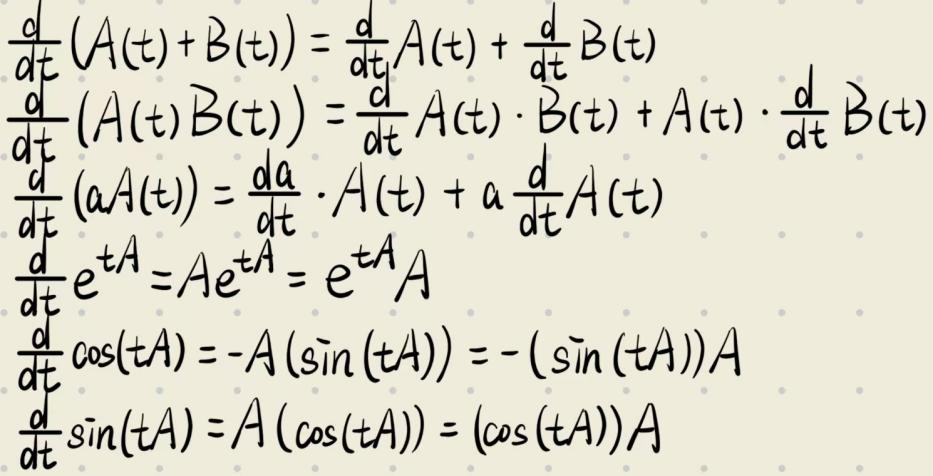


# **2.3.4 函数矩阵对矩阵的导数**

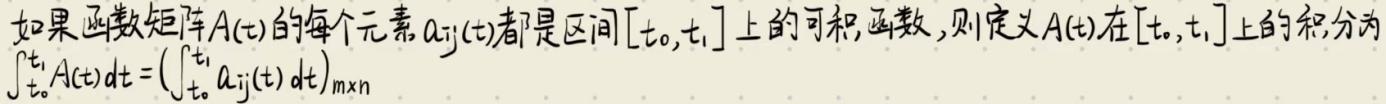
1. 函数矩阵的导数



函数矩阵的导数具有如下性质



1. 函数矩阵的积分



1. 函数对矩阵的导数

设，元函数，定义对矩阵的导数为

1. 函数矩阵对矩阵的导数

设，元函数. 定义函数矩阵

对矩阵的导数为

其中

1. 函数对矩阵的导数的应用举例

函数矩阵对矩阵的导数的应用十分广泛，这里仅举一例——反向传播算法

深度学习中的反向传播算法，即根据损失函数对网络参数的梯度更新网络参数，其核心便是梯度的求解，而这正是函数对矩阵的导数的用武之地

考虑一个层数为的前馈网络，用表示第层神经元的激活值，相邻层间各神经元激活值满足关系

其中表示层间权重，表示偏置值，表示sigmoid函数，其定义公式如下

给定一个输入输出对，定义损失函数. 为了便于计算，定义深度激活值

故有

故

由上可得可以将求损失函数对网络参数的梯度的问题转化求损失函数对深度激活值的导数的问题

若取损失函数

则可求得

故

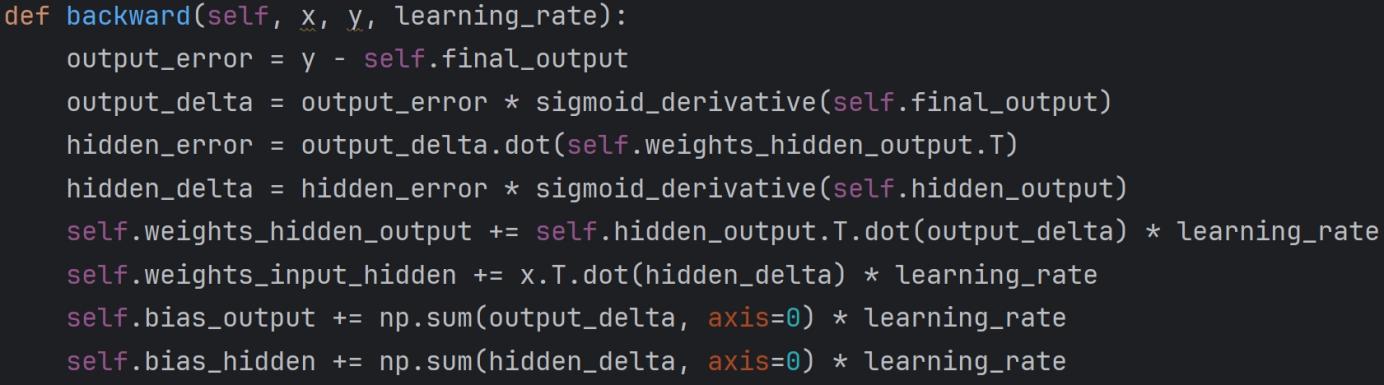
其中是Hadamard积，定义如下

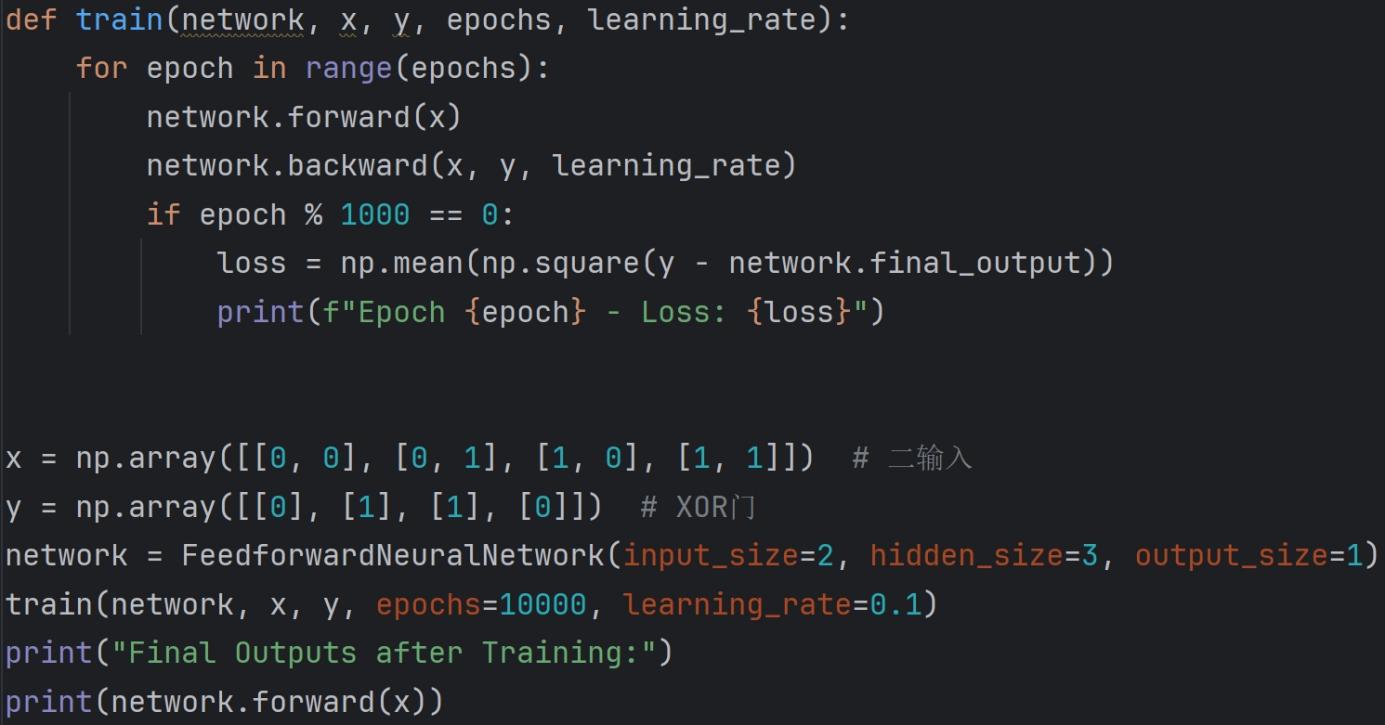
又有

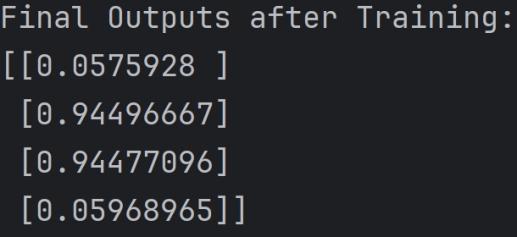
故

由此可顺序地求出每一层的，进而得到与

以下代码实现了一个的前馈网络的反向传播，一定程度上体现了上述思想







在反向传播算法的推导过程中，函数对矩阵的导数将复杂的代数式化为了更简洁的形式，从而本文能更容易观察到其中的规律，找到问题的关键，进而给出准确且高效的解决方案。

# **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**

设阶矩阵的特征多项式为. 如果首多项式

满足且整除（矩阵的最小多项式与特征多项式均满足这些条件）. 那么，的零点都是的特征值。记的互异零点为，相应的重数为，则有

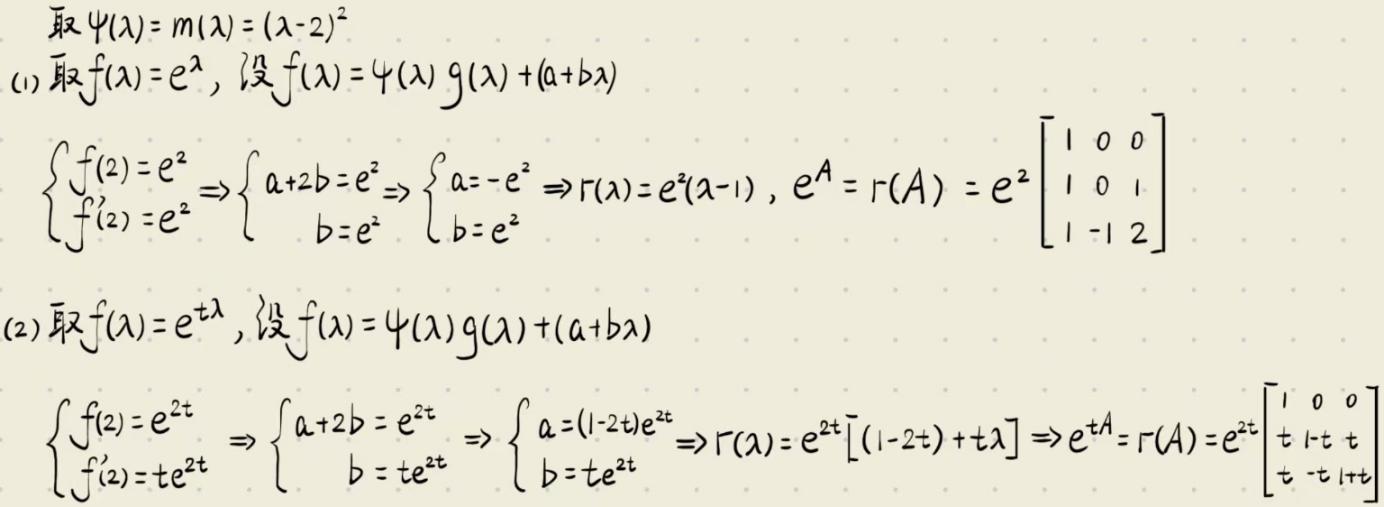
这里，表示的阶导数。设

其中是次数低于的多项式，于是可由

确定出. 利用**，**可得

# **3.1.2 举例展示求法**

设，求与



##### **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**

设首多项式

满足，即

或者

由此可以求出

于是有

这表明，利用上式可以将一个矩阵幂级数的求和问题，转化为个数项级数的求和问题

# **3.2.2 举例展示求法**

设**，**求

. 由于，所以，，，. 于是有

##### **3.3 对角型法**

# **3.3.1 对角型法求矩阵函数的步骤推导**

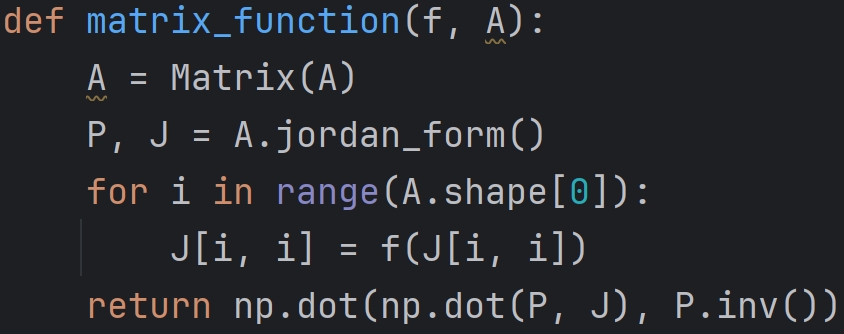
设相似于对角矩阵**，**即有可逆矩阵，使得

则有**，**于是可得

从而

这表明，当与对角矩阵相似时，可以将矩阵幂级数的求和问题转化为求相似变换矩阵的问题

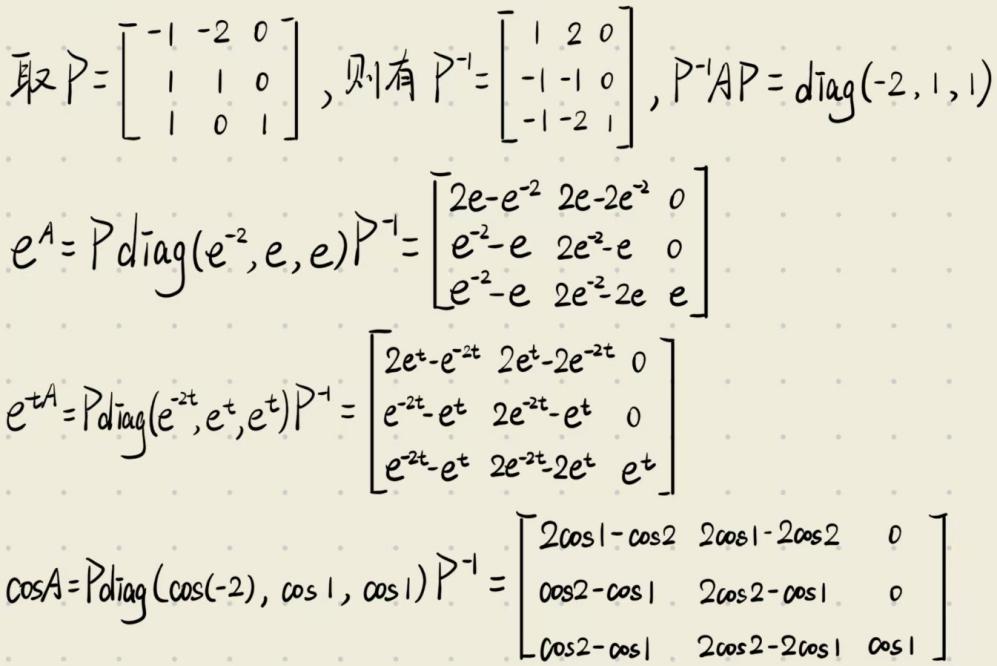
该方法是易于编程实现的，相关代码如下



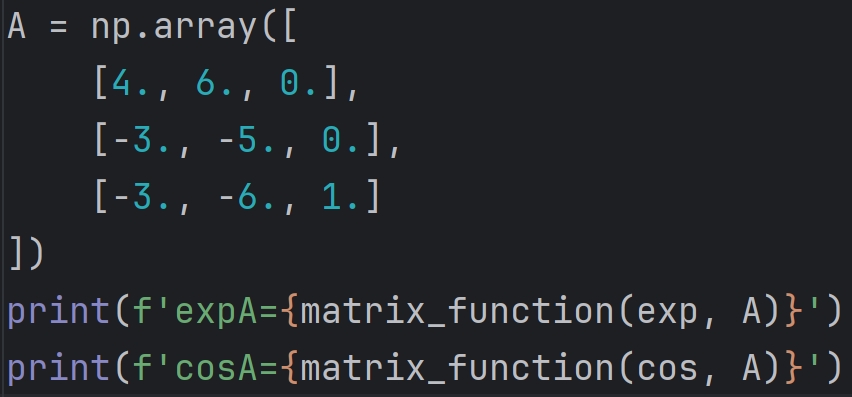
分析可知，该算法的时间复杂度为，其中为的阶数，为使用幂级数近似求时的迭代次数；若直接用定义求**，**复杂度将为，这表明当与对角矩阵相似时，求解矩阵函数的速度可以达到极大的提高

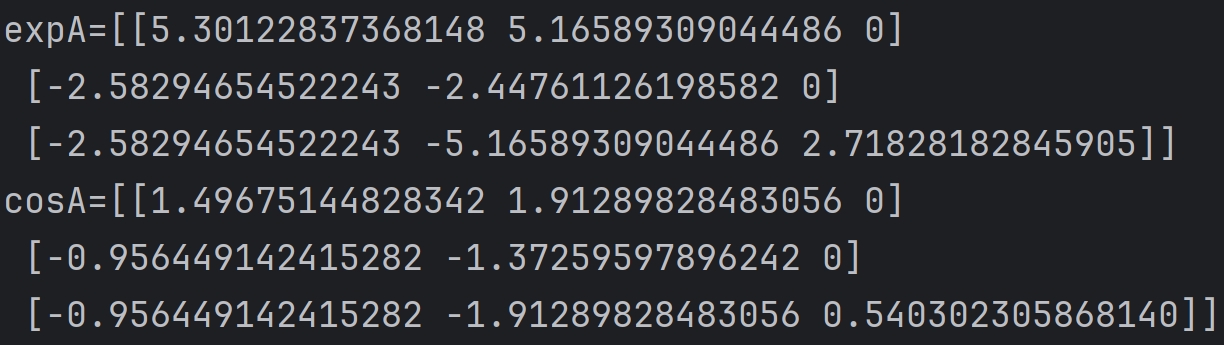
# **3.3.2 举例展示求法**

设**，**分别求及

****

编程检验结果如下





输出结果与理论计算结果一致

##### **3.4 Jordan标准形法**

# **3.4.1 Jordan标准形法求矩阵函数的步骤推导**

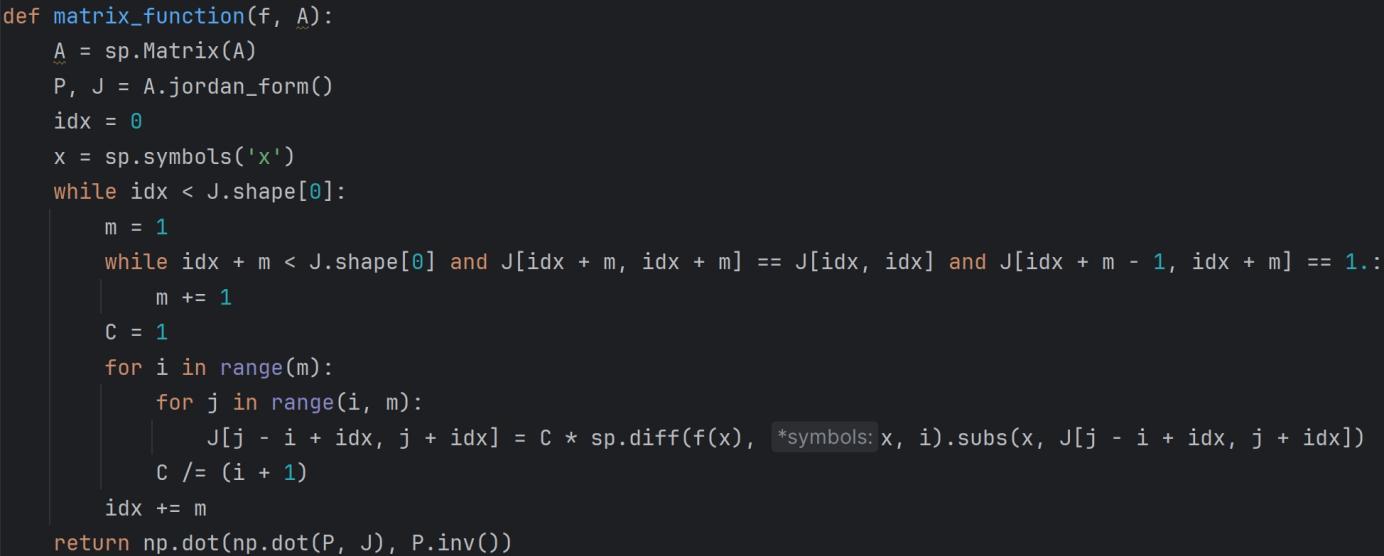
设的Jordan标准形为**，**则有可逆矩阵**，**使得

其中

可求得

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

该方法是易于编程实现的，相关代码如下



分析可知，该算法的时间复杂度为，其中为的阶数，为使用幂级数近似求时的迭代次数；若直接用定义求或是将展开为Jordan标准形后用定义求**，**复杂度均为，故该方法是对矩阵函数求解方法的极大的优化

# **3.4.2 举例展示求法**

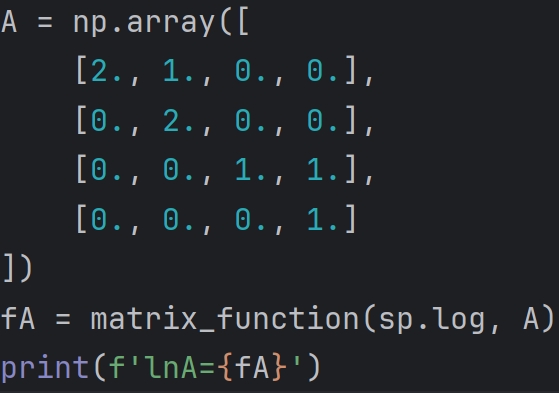
设，求

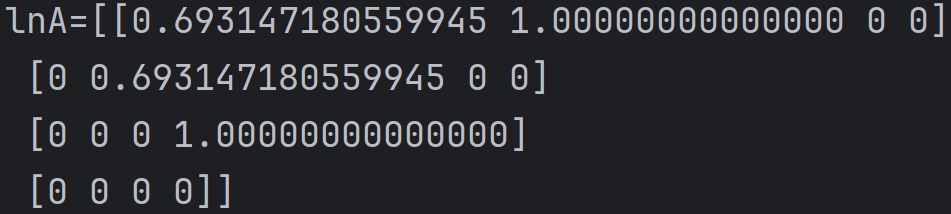
已为Jordan标准形，Jordan块为

分别求出函数值得到

故

编程检验结果如下





输出结果与理论计算结果一致

# **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**

设方阵的任意阶顺序主子式均不为，即

现用Gauss消去法将化为上三角矩阵

构造矩阵序列

其中

是Frobenius矩阵，其中

因为倍加变换不改变矩阵的行列式的值，所以

即

故上述构造成立

计算得到

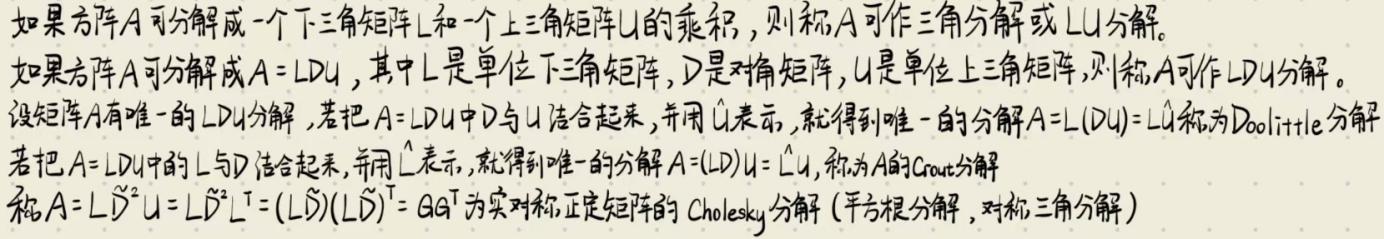
这是上三角矩阵

移项得到

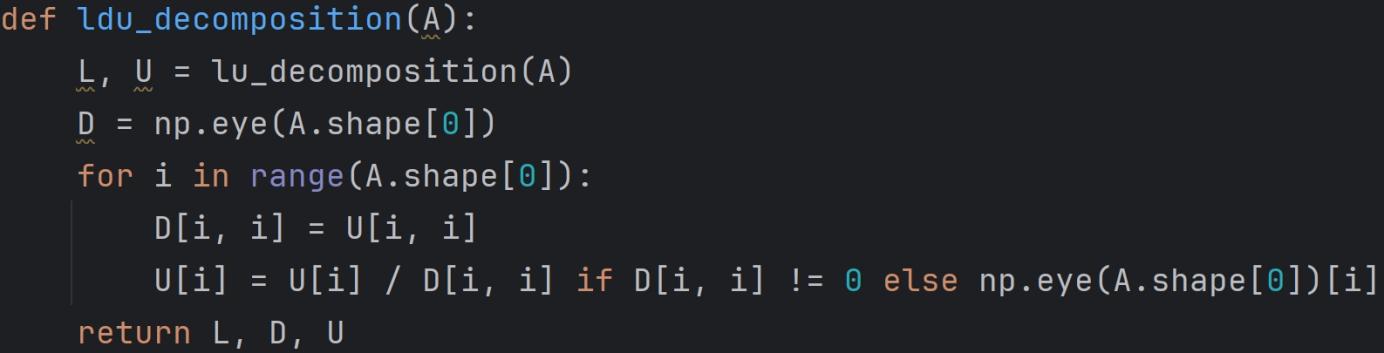
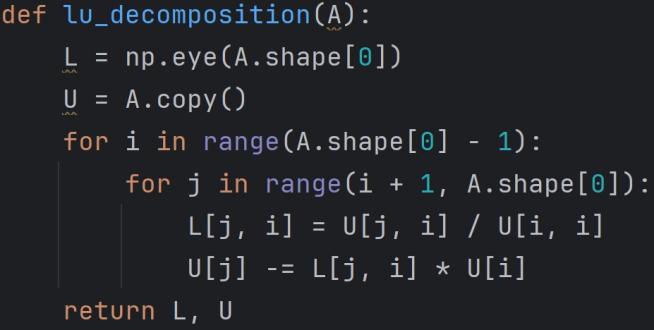
容易求出

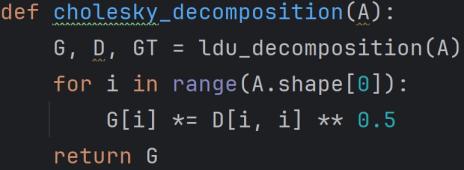
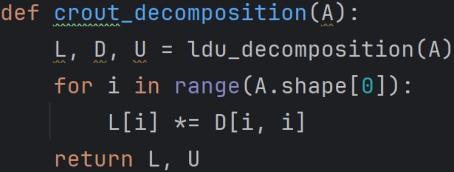
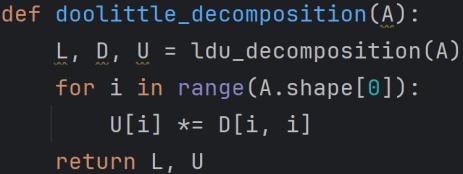
这是单位下三角矩阵. 若令**，**则得

这样就被分解成一个单位下三角矩阵与一个上三角矩阵的乘积



该方法是易于编程实现的，相关代码如下





分析可知，上述所有算法的时间复杂度均为

对于一般的方阵**，**条件

一般很难满足，这导致在分解过程中，条件

可能不成立。但只要，那么元右侧一定存在非零元，设非零元所在列数为，即

其中. 此时令右乘列交换矩阵**，**即

即可使条件成立

记第r次迭代时需要右乘的交换矩阵为**（**若不需要则**），**则递推式应修正为

计算得到

移项得到

记**，**则有

上式称为方阵的LUP分解

对上式稍加变形得

这表明对任意可逆方阵**，**总存在一种排列方式，对的列进行重排，使得重排后的方阵满足

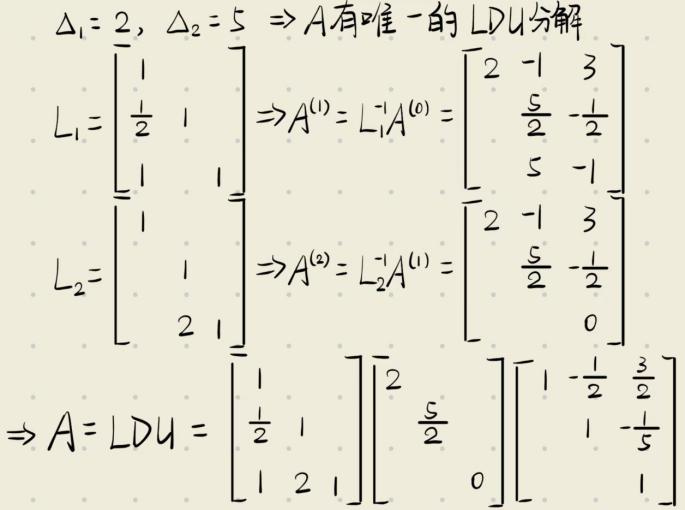
由对称性可得，对任意可逆方阵**，**总存在一种排列方式，对的行进行重排，使得重排后的方阵满足

用矩阵语言描述便是，对任意可逆方阵**，**总存在交换矩阵**，**使得

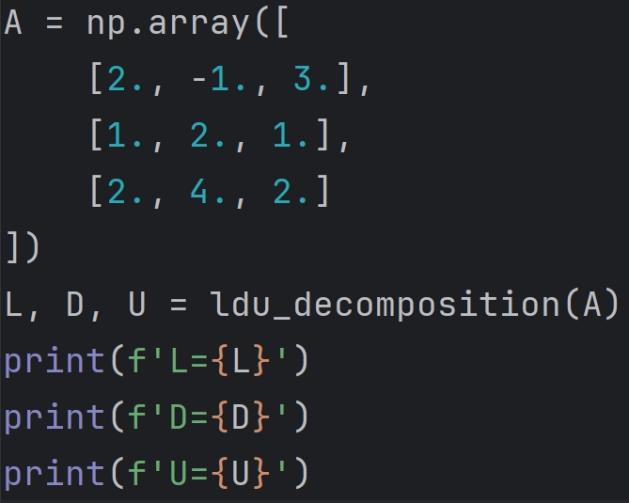
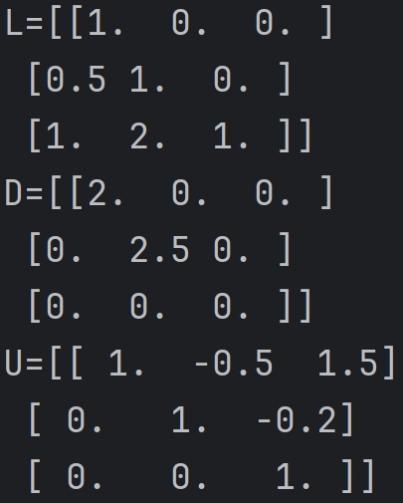
上式称为方阵的PLU分解

# **4.1.2 举例展示求法**

求矩阵的LDU分解，其中



编程检验结果如下

输出结果与理论计算结果一致

矩阵的LU分解的一个常见的应用是解线性方程组

考虑线性方程组

其系数矩阵为方阵。若，则可作PLU分解，即

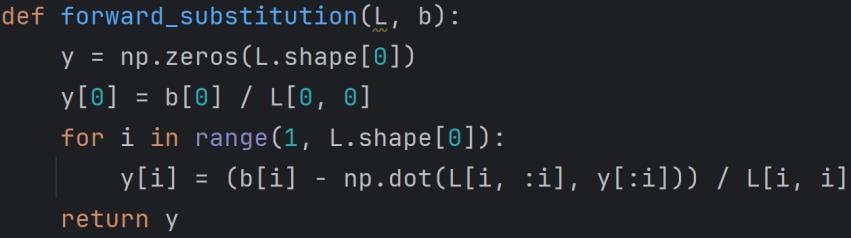
故方程可写为

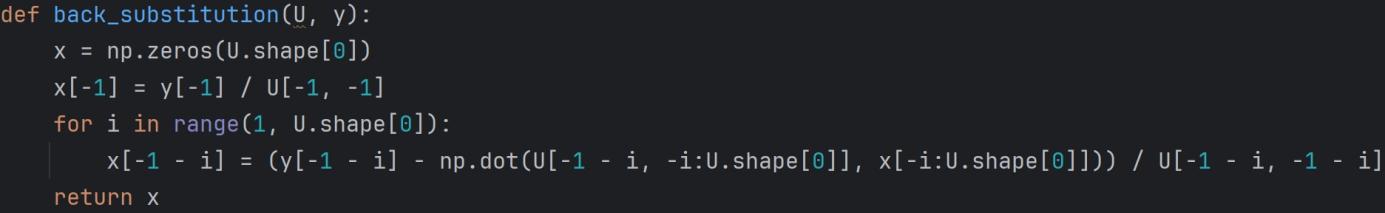
令**，**则可先由

解出**，**再由

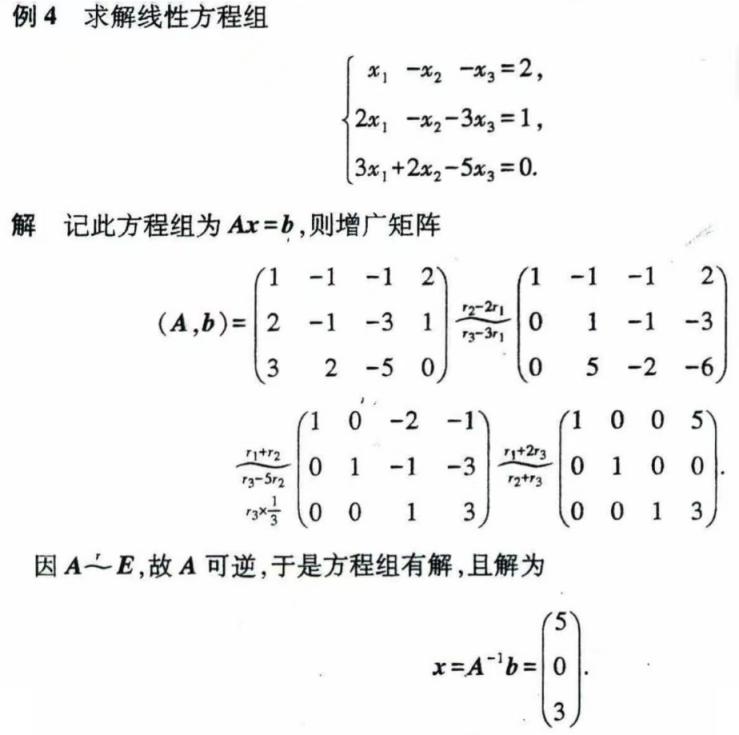
解出**.** 在上述两方程中，系数矩阵均为三角矩阵，故可以通过前向和后向代入直接求解，而无需使用高斯消元

该方法是易于编程实现的，相关代码如下





以同济七版《线性代数》第三章例为例测试如下



4.1.2.8

输出结果与理论计算结果一致

##### **4.2 矩阵的QR分解**

# **4.2.1 矩阵QR分解的步骤推导**

设是实（复）矩阵，且其个列向量线性无关。将它们按Schmidt正交化方法正交化，可得

其中

将上式改写为

用矩阵形式表示为

其中

再对单位化，可得

于是有

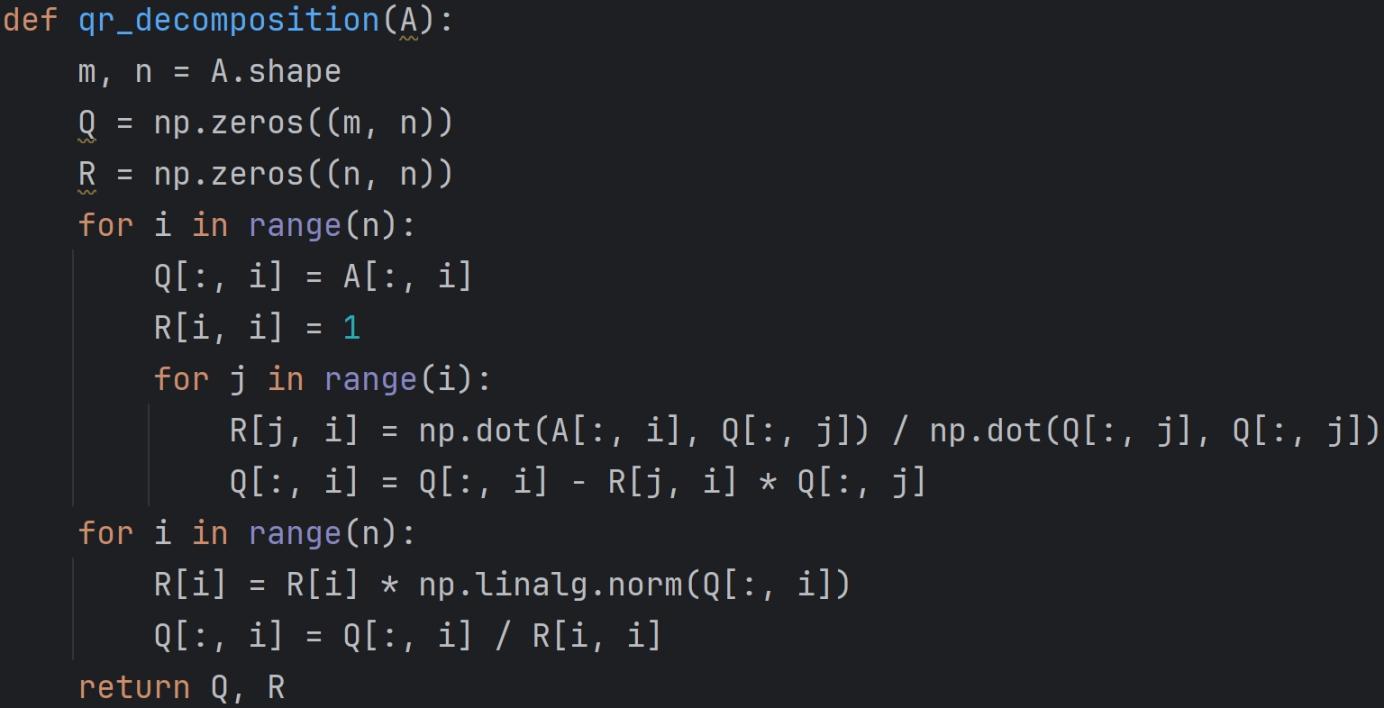
令

则有

其中是实（复）矩阵，且满足，是阶实（复）可逆上三角矩阵。当时，是正交（酉）矩阵。

4.2.1.1

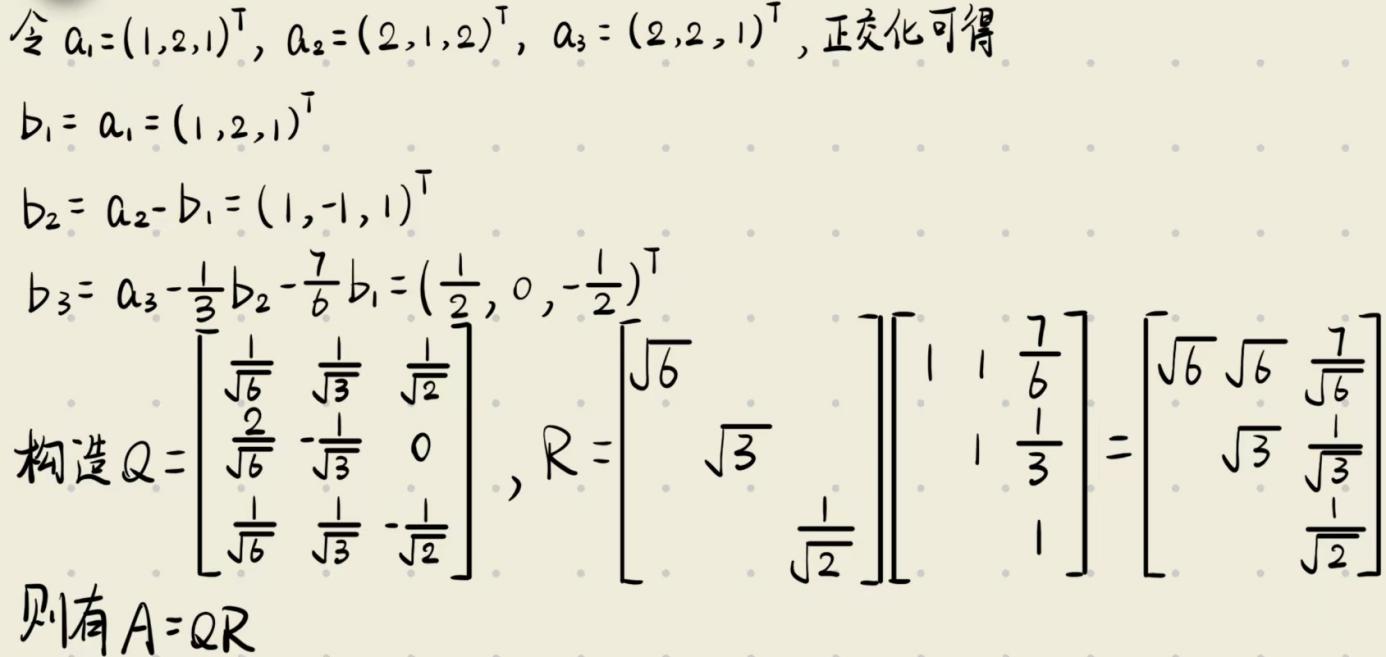
该方法是易于编程实现的，相关代码如下



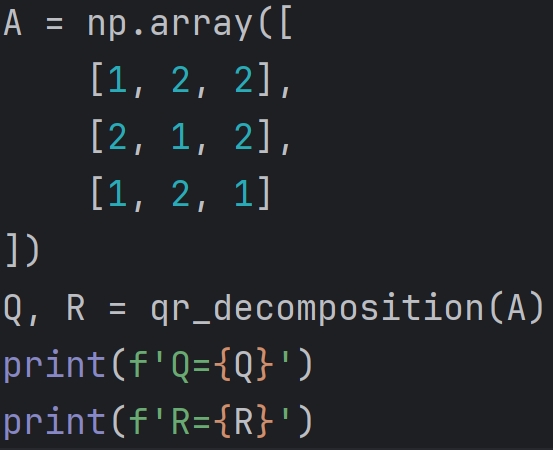
分析可知，该算法时间复杂度为

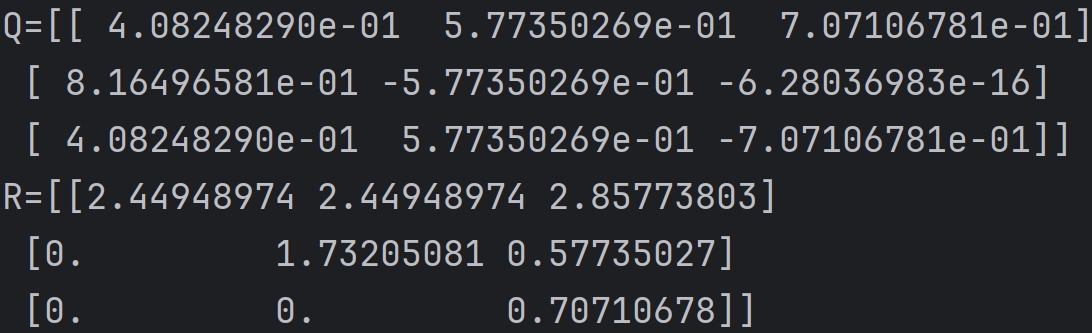
# **4.2.2 举例展示求法**

用Schmidt正交化方法求矩阵的QR分解，其中



编程检验结果如下





输出结果与理论计算结果一致

矩阵的QR分解也可以用于解线性方程组。与LU分解不同的是，QR分解不要求系数矩阵为方阵

考虑线性方程组

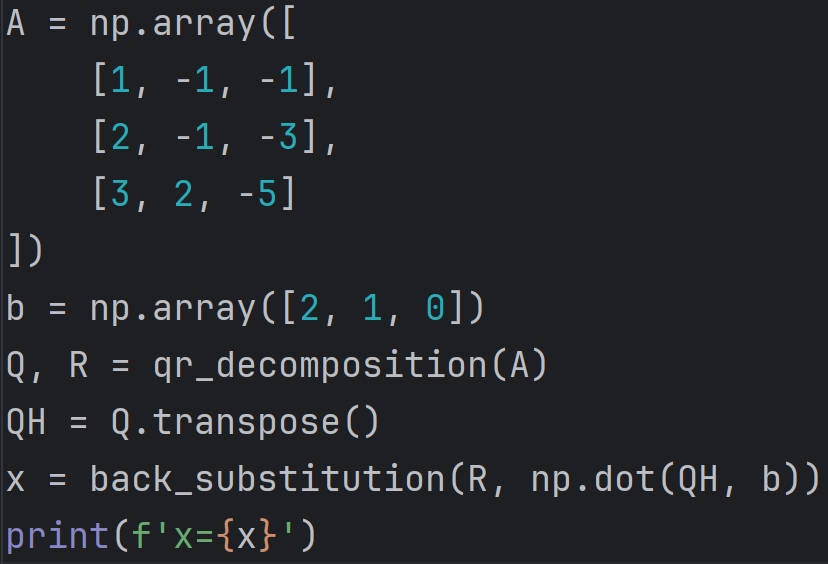
其系数矩阵，且其个列向量线性无关，则存在，阶可逆上三角矩阵，使得

且有

故方程可写为

该方程系数矩阵为上三角矩阵，故可以通过后向代入直接求解，而无需使用高斯消元

该方法是易于编程实现的，相关代码在LU分解部分已展示，依旧以同济七版《线性代数》第三章例为例测试如下（题目见LU分解部分）



4.2.2.5

输出结果与理论计算结果一致

##### **4.3 矩阵的满秩分解**

# **4.3.1 矩阵满秩分解的步骤推导**

设**，**根据矩阵的初等变换理论，对进行初等行变换，可将化为行阶梯型矩阵**，**即

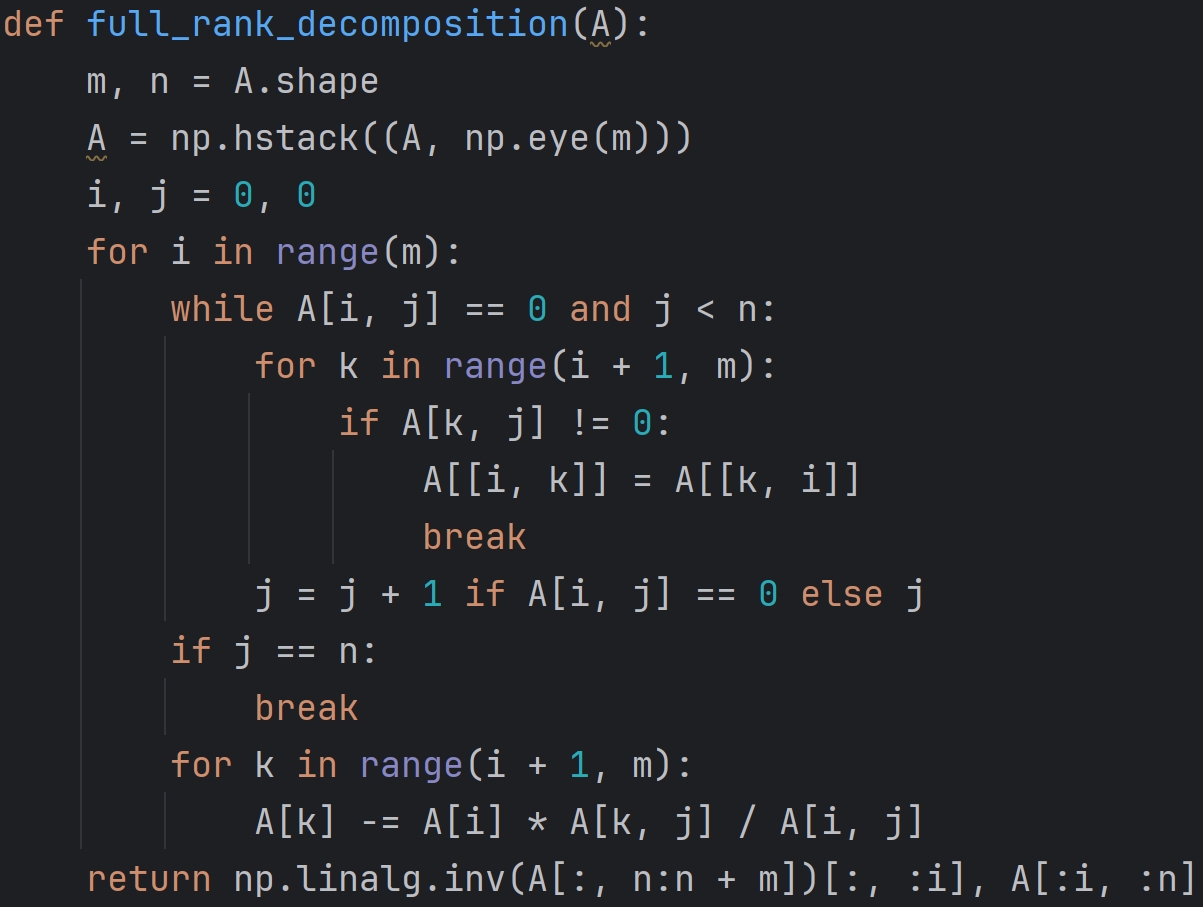
于是存在有限个阶初等矩阵的乘积，记作**，**使得**，**或者.将分块为

则有

其中是列满秩矩阵，是行满秩矩阵

4.3.1.1

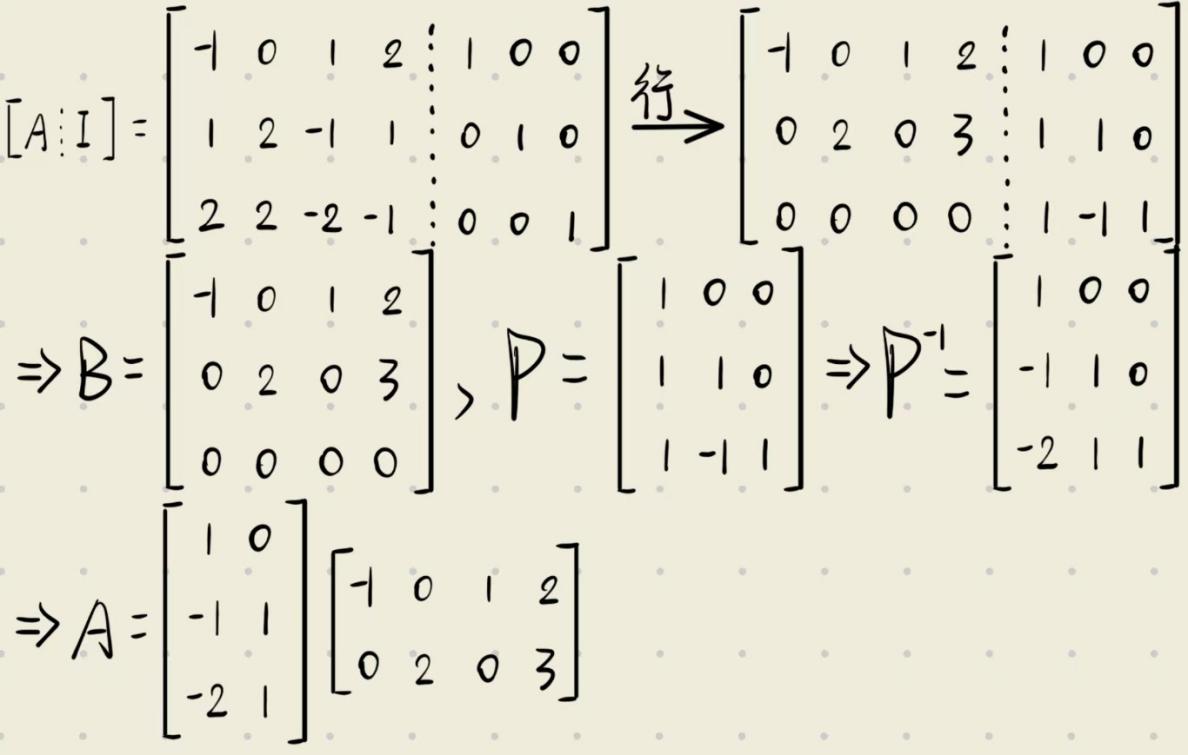
该方法是易于编程实现的，相关代码如下



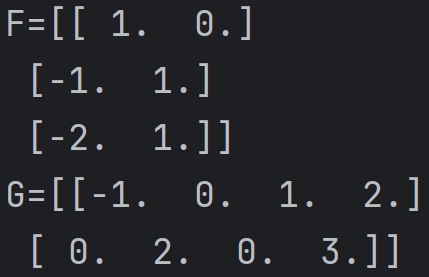
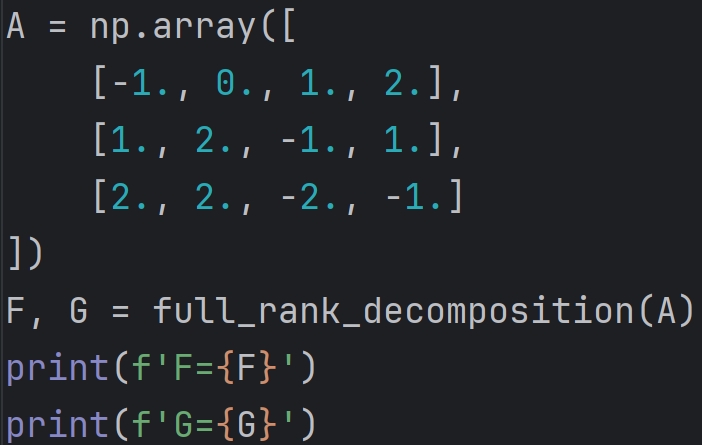
分析可知，该算法时间复杂度为

# **4.3.2 举例展示求法**

求矩阵的满秩分解，其中



编程检验结果如下



输出结果与理论计算结果一致

##### **4.4 矩阵的奇异值分解**

# **4.4.1 矩阵奇异值分解的步骤推导**

本部分将先介绍奇异值的概念，随后介绍矩阵奇异值分解的概念，进而给出当矩阵为实可逆矩阵时的特殊讨论

4.4.1.1

这里是Hermite矩阵，故存在阶酉矩阵，使得

即

其中，而为矩阵的全部非零奇异值。将分块为

则有

令**，**则**，**即的个列是两两正交的单位向量，记作. 将扩充为的标准正交基，记增添的向量为，并构造矩阵**，**则

是阶酉矩阵，且有

于是可得

即

称上式为矩阵的奇异值分解（SVD分解）。

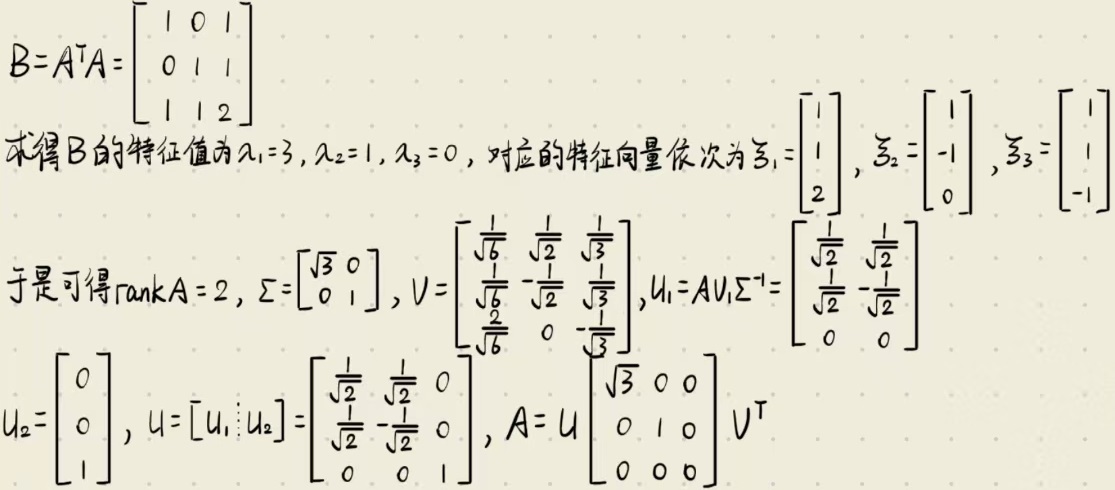
当矩阵为实可逆矩阵时，为实对称矩阵，故存在正交矩阵满足上述矩阵所有性质，同理可用相同的构造方式构造正交矩阵满足上述矩阵的所有性质。因此，当矩阵为实可逆矩阵时，存在正交矩阵和，使得

其中为矩阵的全部奇异值。称上式为矩阵的正交对角分解。

矩阵的正交对称分解理论表明，欧氏空间中的任意可逆的线性变换都可以表示为正交变换和缩放变换的复合，故三维欧氏空间中的任意可逆的线性变换都可以表示为旋转变换和缩放变换的复合，这在3D引擎中有重要应用。3D引擎中的物体可以通过平移，旋转，缩放等变换进行操作。其中，旋转变换和缩放变换是两种基本的线性变换。通过正交对角分解，本文可以将复杂的3D变换分解为更简单的旋转变换和缩放变换，从而更方便地实现和控制物体的变换。

# **4.4.2 举例展示求法**

# 求矩阵的奇异值分解



矩阵的奇异值分解的一个常见的应用是图像压缩

假设一幅图像是灰度图像，可以将其表示为一个矩阵，其中矩阵的元素代表图像中第行第列的像素值。将图像矩阵进行奇异值分解，得到

保留的前行前列，即前大的奇异值，并保留和的前列，其余部分舍弃，即

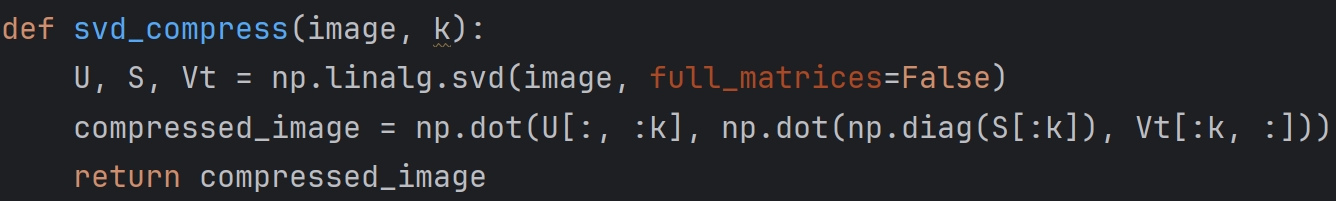
当

时，上式等号右侧存储空间小于左侧，此时可以通过存储等号右侧的矩阵来代替原始矩阵，从而实现图像压缩

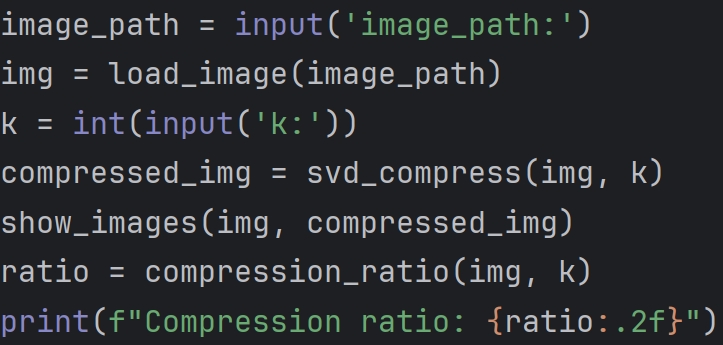
当时，条件变为

即选择的不能超过

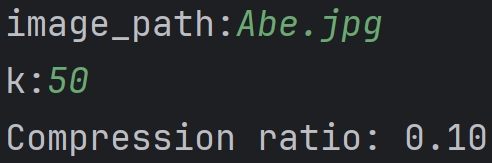
该方法是易于编程实现的，相关代码如下

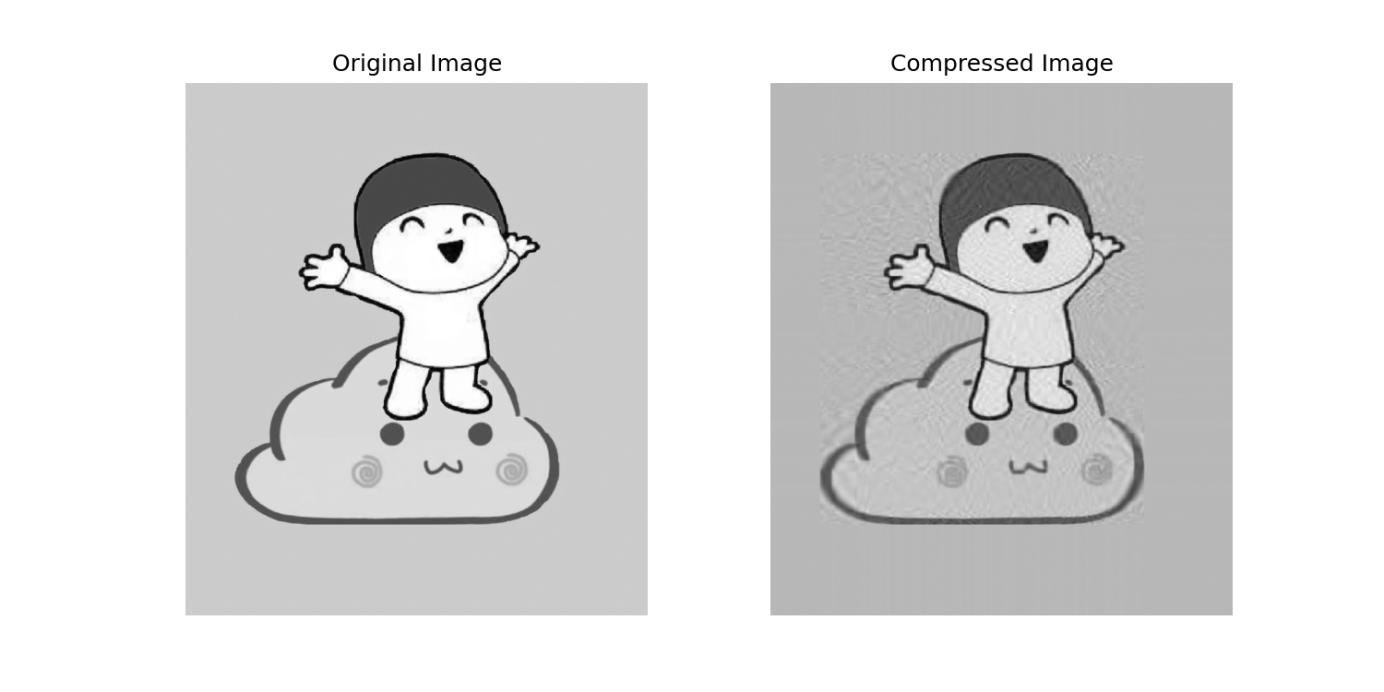


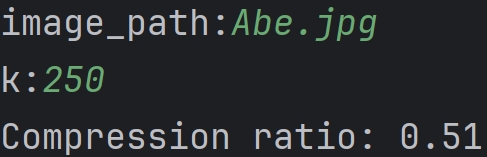
测试压缩两类图像代码如下

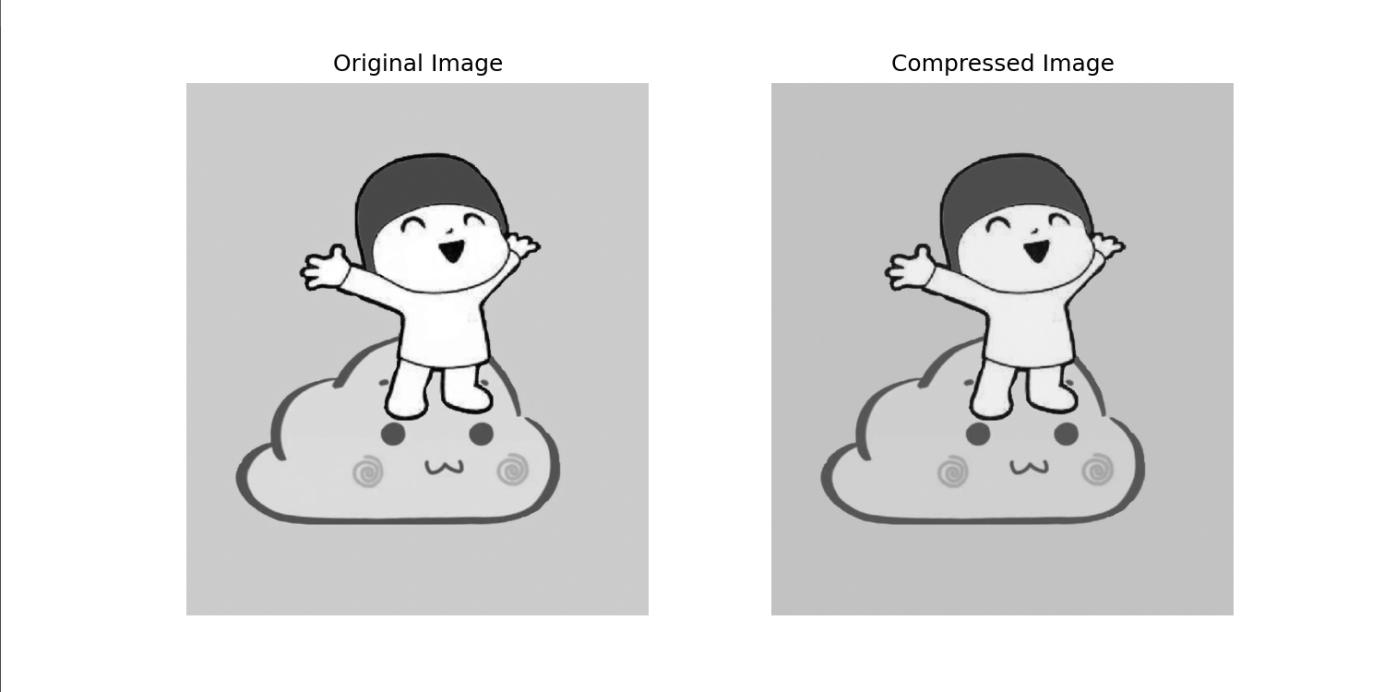


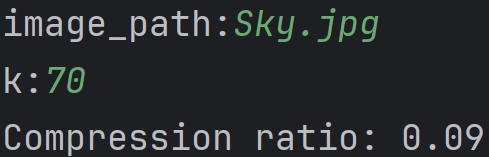
测试结果如下

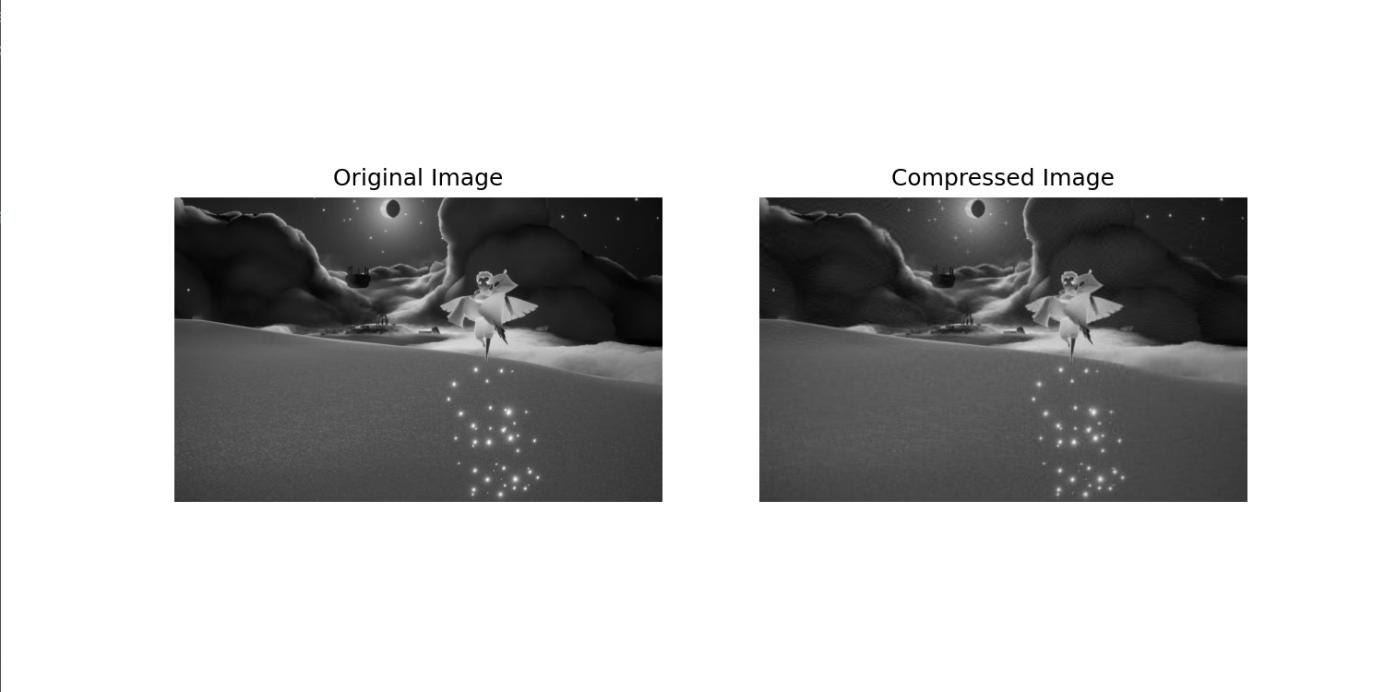


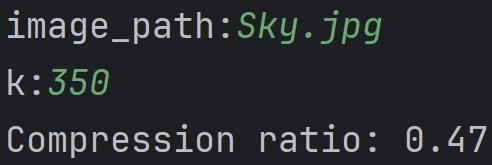


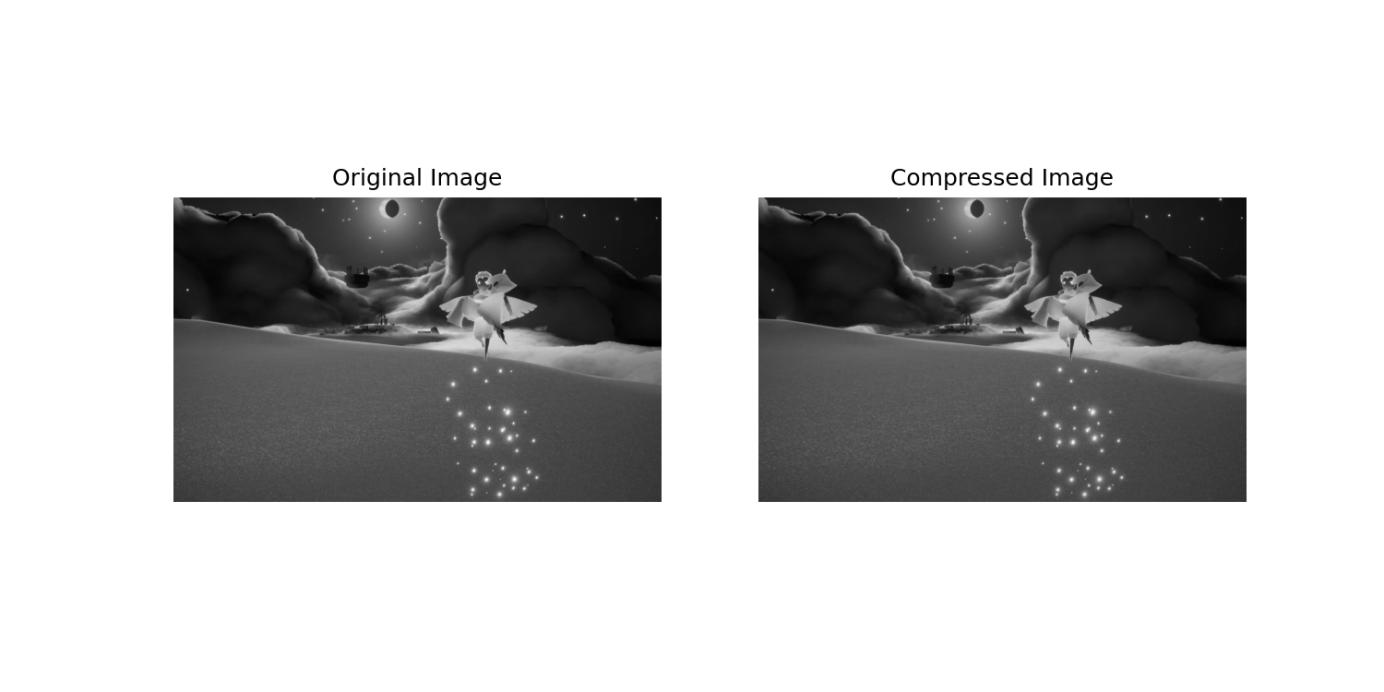












该方法已在各大领域得到应用

1. JPEG图像压缩：JPEG压缩标准就是基于一种类似SVD的技术进行图像压缩。虽然JPEG具体使用的是离散余弦变换（DCT），但它与SVD的原理类似，都是通过分解图像来去除冗余数据。
2. 医学成像：在医学成像中，SVD压缩用于存储和传输高分辨率图像，如CT扫描图像或MRI图像。在这些领域中，图像的大小可能非常庞大，SVD能够有效减少存储需求。
3. 视频压缩：视频图像的每一帧都可以看作一个图像矩阵，通过对每一帧应用SVD，也可以实现视频压缩。

##### **4.5 利用矩阵分解求矩阵广义逆**

# **4.5.1 矩阵广义逆介绍**

本部分将先介绍线性方程组求解问题，随后介绍矩阵广义逆的概念，进而介绍矩阵广义逆在求解线性方程组时的应用。关于矩阵广义逆求解问题将在后续部分介绍

考虑线性方程组

其中. 其解有下列三种情况

1. 有唯一解：此时该解即为所求
2. 有无穷多解：此时求出具有极小范数的解，即

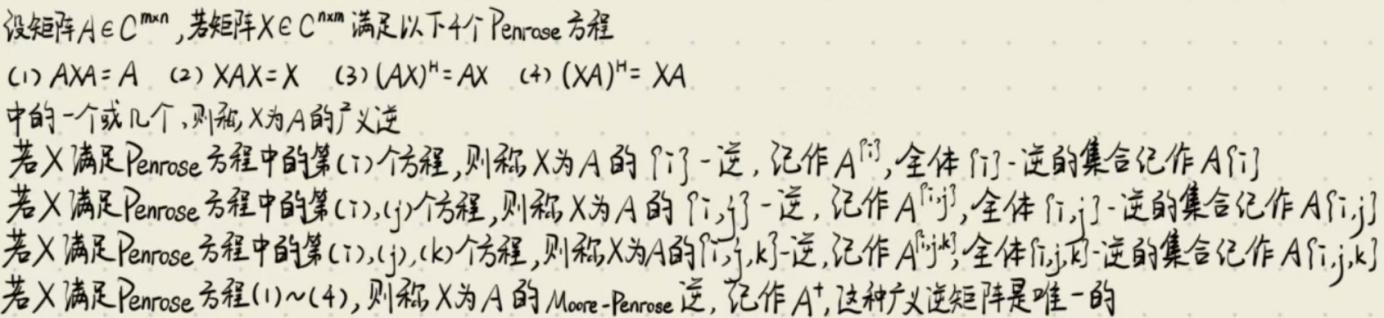
可以证明，满足该条件的解是唯一的，称之为极小范数解

1. 无解：此时求出极值问题

的解，称之为最小二乘解。一般说来，最小二乘解是不唯一的，但在最小二乘解的集合中，具有极小范数的解

是唯一的，称之为极小范数最小二乘解

矩阵广义逆与线性方程组的求解有着极为密切的关系，下面介绍矩阵广义逆的概念



上述定义中矩阵的Moore-Penrose逆是求解线性方程组的极小范数最小二乘解的关键。事实上，方程的极小范数最小二乘解即为

# **4.5.2 利用矩阵满秩分解求矩阵广义逆**

设矩阵的满秩分解为

令

则有

故

且有

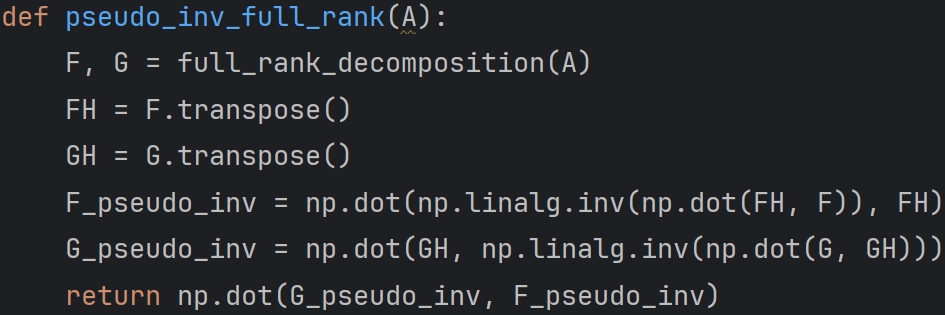
同理可得

令

则有

故

该方法是易于编程实现的，相关代码如下



分析可知，该算法时间复杂度为

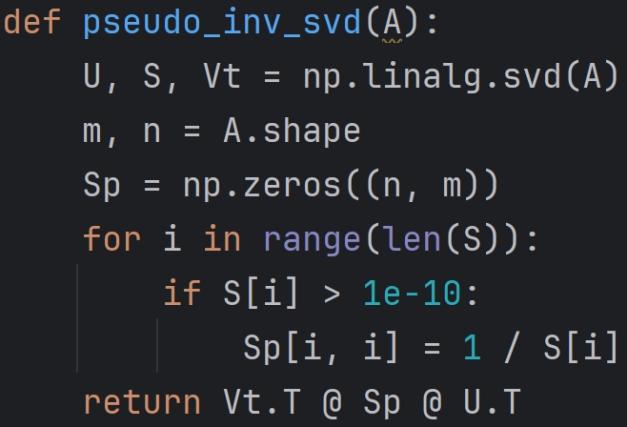
# **4.5.3 利用矩阵奇异值分解求矩阵广义逆**

设矩阵的奇异值分解为

令

则满足个Penrose方程，故有

该方法是易于编程实现的，相关代码如下

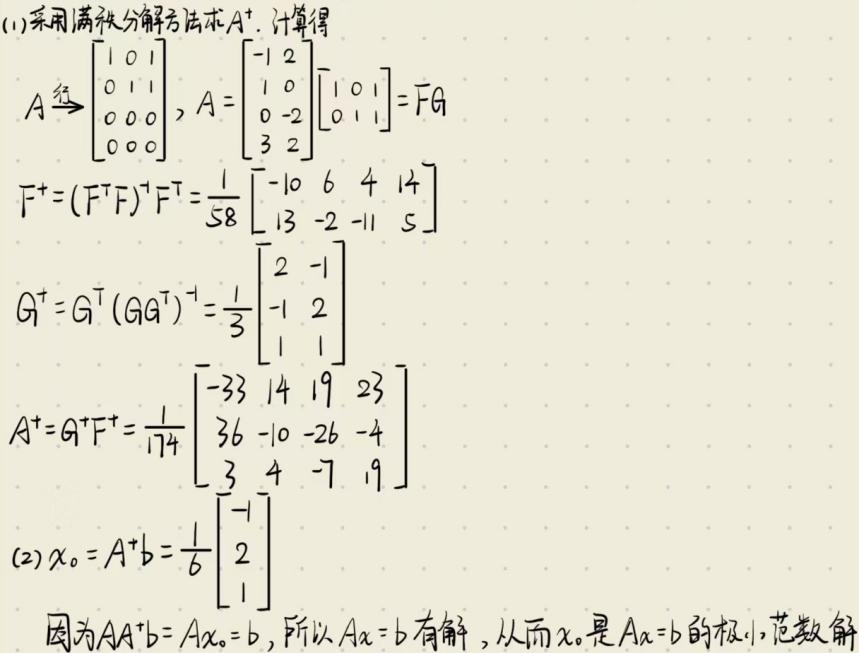


分析可知，该算法时间复杂度为

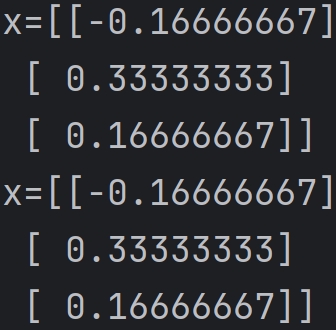
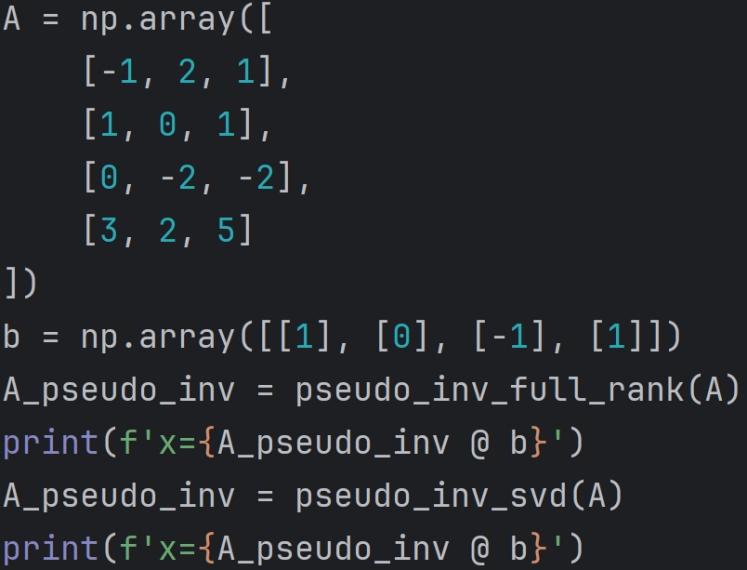
# **4.5.4 举例展示求法**

已知.

1. 求的Moore-Penrose逆
2. 用广义逆矩阵方法判断线性方程组是否有解，并求其极小范数解或者极小范数最小二乘解



编程检验结果如下



输出结果与理论计算结果一致

不相容线性方程组求解理论广泛应用与许多领域，这里仅举一例——回归分析

讨论下述线性回归模型

其中都是与无关的未知参数

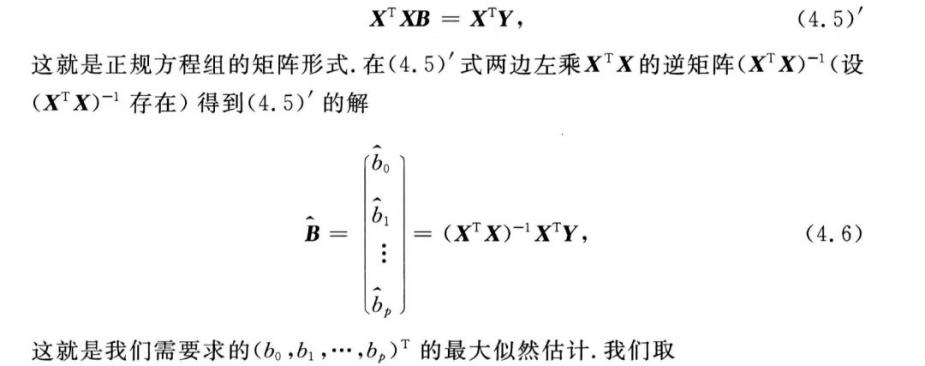
设

是一个样本，引入矩阵

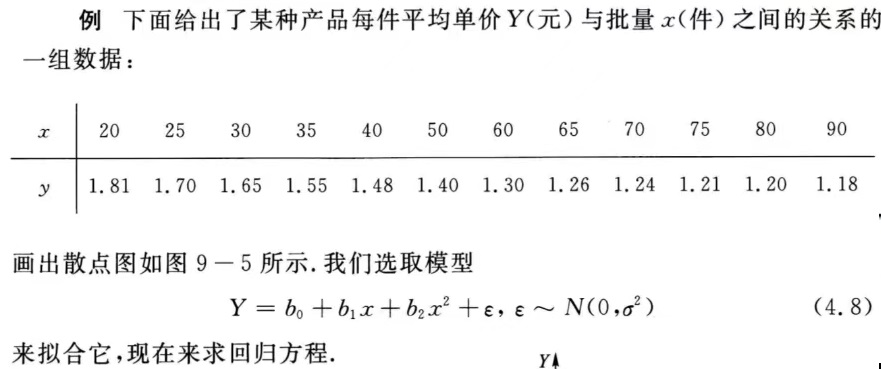
则有

解得

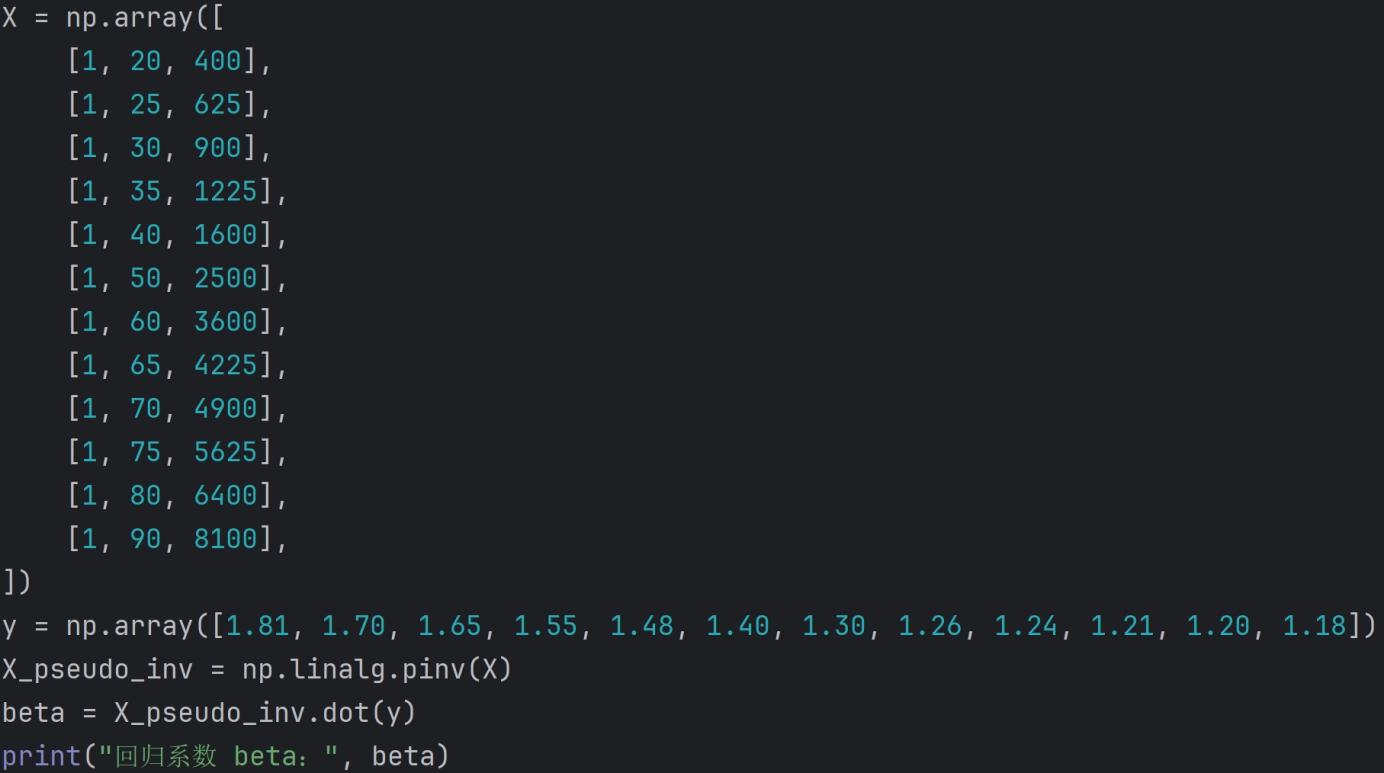
这就是我们需要求的的估计，该结论与浙大五版《概率论与数理统计》给出的结论一致，且基于不列满秩的情况做了推广



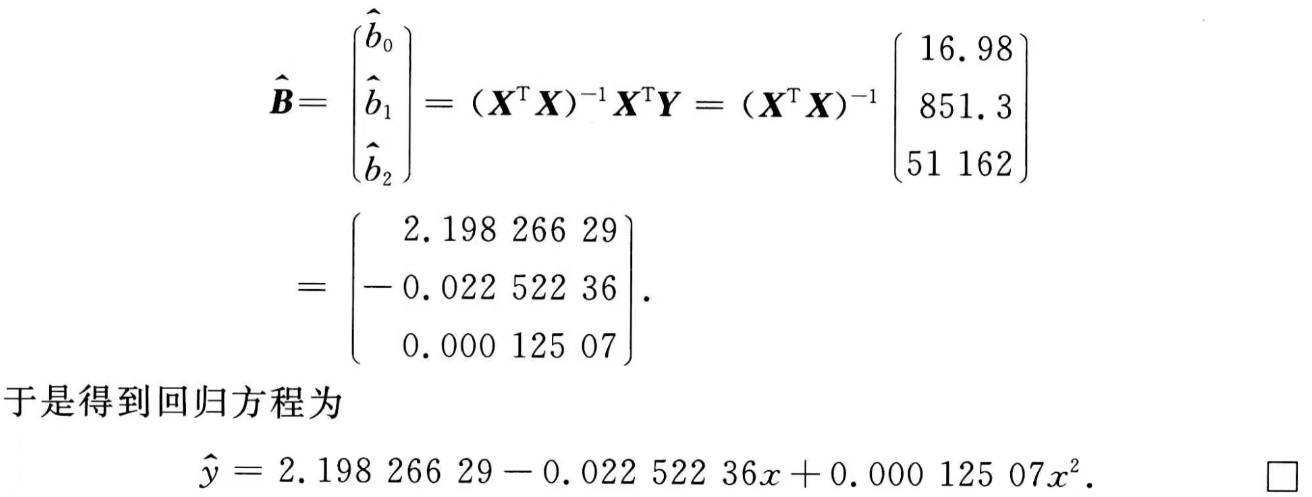
该方法是易于编程实现的，现以浙大五版《概率论与数理统计》P258例题测试



相关代码即输出结果如下



微信截图_20241227124628



输出结果与理论计算结果一致

# **第五章 总结**

本文系统地研究了矩阵论中的核心概念和方法，包括线性空间、线性变换、Jordan标准形、向量范数、矩阵范数、矩阵函数、以及矩阵分解等。通过深入的理论分析和算法推导，我们不仅掌握了这些概念的数学本质，还将其应用于多个实际问题中。

本文的主要闪光点在于：

1. 理论深度与广度

本文对线性空间、Jordan 标准形等概念进行了深入探讨，并结合具体案例分析了其在不同领域的应用。

* 1. 线性空间概念的拓展与应用：

本文从线性空间的基本概念出发，通过多项式空间和矩阵空间的例子，深入探讨了线性空间概念的广泛性和灵活性，揭示了其在数值分析、计算机图形学、机器学习等多个领域的关键作用。这为后续讨论奠定了理论基础，展示了线性空间理论在现代科学技术中的重要地位。

* 1. Jordan标准形的详细求解与意义：

本文详细介绍了 Jordan 标准形的概念、求解方法及其重要性。通过多项式矩阵理论，本文系统地阐述了Jordan标准形的求解过程，并强调了其在简化矩阵问题处理方面的关键作用。该部分为后续分析矩阵函数奠定了基础。

* 1. 欧氏空间中线性变换的深入分析：

本文深入探讨了欧氏空间中线性变换的性质，通过一个详细的例题，将理论一般化，突出了对称变换的可对角化性，并将其扩展到酉空间中的 Hermite 变换。这为我们理解线性变换在欧氏空间中的行为提供了深入的见解。

* 1. 范数理论及其应用：

本文深入探讨了向量范数和矩阵范数的相关概念。通过范数等价性，及其在K-means聚类算法中的应用案例，表明范数在机器学习算法中起着重要作用。此外，我们还讨论了矩阵的可逆性条件、条件数和谱半径等概念，完善了矩阵分析的理论体系。

1. 算法优化与实现

本文系统地介绍了多种求解矩阵函数和矩阵分解的方法，并通过代码实现和时间复杂度分析，展示了优化算法的有效性。

* 1. 多种矩阵函数求解方法的对比分析：

本文深入探讨了多种求解矩阵函数的方法，包括待定系数法、数项级数求和法、对角形法和 Jordan 标准形法。通过对比分析，本文强调了对角形法和 Jordan 标准形法在计算效率上的优势，并提供了代码实现，证明了这些方法的有效性和高效性。

* 1. 矩阵分解方法的详细推导与代码实现：

本文详细介绍了矩阵的 LU 分解（包括 LUP 和 PLU 分解）、QR 分解、满秩分解和奇异值分解 (SVD) 方法，并给出了相应的代码实现。这些分解方法不仅为矩阵理论提供了重要的工具，而且在解决实际问题中具有重要的应用价值，如求解线性方程组和进行图像压缩。

* 1. 利用矩阵分解求广义逆：

本文探讨了矩阵广义逆的概念及其在求解线性方程组中的应用。我们提出了利用满秩分解和奇异值分解来求解矩阵广义逆的方法，并提供了相应的代码实现，证明了矩阵分解在求解广义逆中的效率。

1. 理论联系实际

本文将抽象的矩阵论知识与计算机图形学、机器学习、图像压缩等具体应用相结合，展示了矩阵论在现代科学技术中的重要作用。

* 1. 线性变换在计算机图形学中的应用：明确指出旋转变换的理论基础。
  2. 多项式线性空间在数值分析中的应用：强调了多项式空间在函数近似中的作用。
  3. 向量范数在 K-means 聚类算法中的应用：不仅介绍了 K-means 算法，还给出了基于 2-范数的代码实现，并分析了范数在算法中的作用。
  4. 函数对矩阵导数在反向传播算法中的应用：详细推导了反向传播算法中梯度的计算过程，并通过代码实现了一个简单的前馈网络。
  5. 矩阵的正交对角分解在 3D 引擎中的应用：揭示了正交对角分解在简化 3D 变换中的作用。
  6. 矩阵的奇异值分解在图像压缩中的应用：通过代码实现演示了 SVD 在图像压缩中的实际效果，并分析了其在不同领域的应用。
  7. 矩阵广义逆在求解线性方程组中的应用：详细阐述了广义逆在求解不同类型线性方程组中的作用。
  8. 不相容方程组求解在回归分析中的应用：详细阐述了不相容方程组求解在回归分析中的应用，对浙大五版《概率论与数理统计》给出的结论做出了推广，同时给出了代码实现。

通过本研究，我们不仅加深了对矩阵论的理解，也提高了解决实际问题的能力。本文的研究成果为相关领域的研究人员提供了一定的理论参考和实践指导。

##### 参考文献

[1] 张凯院, 徐仲等. 矩阵论[M]. 西北工业大学出版社. 2017年8月.