数值代数— 绪论

邓建松

2018 年 9 月 11 日

数据处理问题

- 现代的科学技术发展十分迅速，它们有一个共同的特点，就是都有大量的数据问题

数据处理问题

- 现代的科学技术发展十分迅速，它们有一个共同的特点，就是都有大量的数据问题

  1. 发射一颗卫星，从卫星设计开始到发射、回收为止，科学家和工程技术人员、工人就要对卫星的总体部件进行全面的设计和生产，要对选用的火箭进行设计和生产

数据处理问题

- 现代的科学技术发展十分迅速，它们有一个共同的特点，就是都有大量的数据问题
  1. 发射一颗卫星，从卫星设计开始到发射、回收为止，科学家和工程技术人员、工人就要对卫星的总体部件进行全面的设计和生产，要对选用的火箭进行设计和生产
  2. 在高能加速器里进行高能物理试验，研究具有很高能量的基本粒子的性质、它们之间的相互作用和转化规律，这里面也有大量的数据计算问题

科学活动的新手段

- 1946年第一台电子计算机诞生

科学活动的新手段

- 1946年第一台电子计算机诞生
- 理论研究、科学试验、科学计算为当今世界科学活动的三种主要方式

科学活动的新手段

- 1946年第一台电子计算机诞生
- 理论研究、科学试验、科学计算为当今世界科学活动的三种主要方式
- 为科学与工程问题提供计算方法，提高计算的可靠性、有效性和精确性，是科学与工程计算领域的主要研究内容

科学活动的新手段

数值代数基本问题

研究的必要性

主要技巧

误差分析

算法速度

考核方式

- 1946年第一台电子计算机诞生
- 理论研究、科学试验、科学计算为当今世界科学活动的三种主要方式
- 为科学与工程问题提供计算方法，提高计算的可靠性、有效性和精确性，是科学与工程计算领域的主要研究内容
- 研究计算问题的解决方法和有关数学理论问题的一门学科就叫做计算数学

科学活动的新手段

- 1946年第一台电子计算机诞生
- 理论研究、科学试验、科学计算为当今世界科学活动的三种主要方式
- 为科学与工程问题提供计算方法，提高计算的可靠性、有效性和精确性，是科学与工程计算领域的主要研究内容
- 研究计算问题的解决方法和有关数学理论问题的一门学科就叫做计算数学
- 计算数学属于应用数学的范畴

数值线性代数

- 简称为数值代数，也称为矩阵计算

数值线性代数

数值代数基本问题
研究的必要性
主要技巧
误差分析
算法速度
考核方式

- 简称为数值代数，也称为矩阵计算
- 它是科学与工程计算的核心

数值线性代数

- 简称为数值代数，也称为矩阵计算
- 它是科学与工程计算的核心
    - 大部分科学与工程计算问题最终都归结为一个矩阵计算问题，特别是大规模矩阵计算问题

数值线性代数

数值代数— 绪论
邓建松
数值代数基本问题
研究的必要性
主要技巧
误差分析
算法速度
考核方式

- 简称为数值代数，也称为矩阵计算
- 它是科学与工程计算的核心
  - 大部分科学与工程计算问题最终都归结为一个矩阵计算问题，特别是大规模矩阵计算问题
- 数值代数的研究内容就是针对各类科学与工程问题的特点，设计出相应的快速可靠的算法

“计算数学”专业课程

- 本科阶段

“计算数学”专业课程

数值代数— 绪论

邓建松

数值代数基本问题

研究的必要性

主要技巧

误差分析

算法速度

考核方式

- 本科阶段
  - 数值代数、数值分析、偏微分方程数值解、有限元、计算机图形学

“计算数学”专业课程

- 本科阶段
  - 数值代数、数值分析、偏微分方程数值解、有限元、计算机图形学
- 研究生阶段

“计算数学”专业课程

- 本科阶段
  - 数值代数、数值分析、偏微分方程数值解、有限元、计算机图形学
- 研究生阶段
  - 计算流体力学、高级有限元、并行计算及算法、计算机辅助几何设计、样条函数与逼近论、多变量函数逼近论、计算代数几何、小波分析、高级几何建模与图形学

基本问题

- 求解线性方程组：给定n阶非奇异矩阵A和n维向量b, 求解方程$Ax = b$, 其中x是未知的n维向量

基本问题

数值代数— 绪论

邓建松

数值代数基本问题

研究的必要性

主要技巧

误差分析

算法速度

考核方式

- 求解线性方程组：给定n阶非奇异矩阵A和n维向量b, 求解方程$Ax = b$, 其中x是未知的n维向量
- 线性最小二乘问题：给出$m \times n$阶矩阵A和m维向量b, 求n维向量x使得

$$\|Ax - b\|_2 = \min\{\|Ay - b\|_2 : y \in \mathbb{R}^n\}$$

基本问题

- 求解线性方程组：给定n阶非奇异矩阵A和n维向量b，求解方程$Ax = b$，其中x是未知的n维向量
- 线性最小二乘问题：给出$m \times n$阶矩阵A和m维向量b，求n维向量x使得

$$\|Ax - b\|_2 = \min\{\|Ay - b\|_2 : y \in \mathbb{R}^n\}$$

- 矩阵特征值问题：计算给定方阵A的部分或全部特征值与对应的特征向量

基本问题的变体

- 约束最小二乘问题、完全最小二乘问题、矩阵方程的求解、矩阵函数的计算、广义特征值问题、非线性特征值问题、特征值反问题、奇异值分解的计算

基本问题的变体

- 约束最小二乘问题、完全最小二乘问题、矩阵方程的求解、矩阵函数的计算、广义特征值问题、非线性特征值问题、特征值反问题、奇异值分解的计算
- 特别地，奇异值分解的计算有广泛的应用，也有人称其为数值代数的第四大问题

Cramer法则

上述问题的数学理论已相当完善，但是理论上漂亮的结果在实际计算时有可能不可行：先以Cramer法则为例

- Cramer法则：把线性方程组的求解归结为计算$n+1$个n阶行列式

Cramer法则

上述问题的数学理论已相当完善，但是理论上漂亮的结果在实际计算时有可能不可行：先以Cramer法则为例

- Cramer法则：把线性方程组的求解归结为计算$n+1$个n阶行列式
- Laplace展开定理：计算行列式的理论公式—n阶行列式的乘法次数$> n!$

Cramer法则

- 采用Cramer法则求解线性方程组的乘法运算次数大于$(n+1)!$. 例如，求解一个25阶线性方程组，采用此方法，需要约13亿年

Cramer法则

数值代数基本问题

研究的必要性

主要技巧

误差分析

算法速度

考核方式

- 采用Cramer法则求解线性方程组的乘法运算次数大于$(n+1)!$. 例如，求解一个25阶线性方程组，采用此方法，需要约13亿年
- 而采用Gauss消元法，则可以不超过1秒钟完成求解

Jordan分解

- 基于Jordan标准型与分解定理，可以清楚地知道矩阵所有的特征值及有关信息：几何重数、代数重数、对应的特征向量等

Jordan分解

- 基于Jordan标准型与分解定理，可以清楚地知道矩阵所有的特征值及有关信息：几何重数、代数重数、对应的特征向量等
- 而Jordan分解是非常不稳定的，变换矩阵常常是非常病态的

Jordan分解

- 基于Jordan标准型与分解定理，可以清楚地知道矩阵所有的特征值及有关信息：几何重数、代数重数、对应的特征向量等
- 而Jordan分解是非常不稳定的，变换矩阵常常是非常病态的
- 实际计算采用的通常是具有良好数值性态的Schur分解

现状与问题

- 数值代数只有半个多世纪的历史

现状与问题

- 数值代数只有半个多世纪的历史
- 相关的方法和理论已发展得相对成熟

现状与问题

- 数值代数只有半个多世纪的历史
- 相关的方法和理论已发展得相对成熟
- 大规模矩阵计算问题仍是目前的研究核心问题之一

矩阵分解

- 如何根据矩阵计算问题的特点，设计出有效的计算方法，这是我们关注的首要问题

矩阵分解

- 如何根据矩阵计算问题的特点，设计出有效的计算方法，这是我们关注的首要问题
- 基本想法就是将一般的问题转化为一个或几个易于求解的特殊问题

矩阵分解

数值代数基本问题

研究的必要性

主要技巧

误差分析

算法速度

考核方式

- 如何根据矩阵计算问题的特点，设计出有效的计算方法，这是我们关注的首要问题
- 基本想法就是将一般的问题转化为一个或几个易于求解的特殊问题
- 完成这一转化的基本技巧就是矩阵分解

例

- 对于线性方程组求解问题：当系数矩阵A为下三角或者上三角时，方程组的求解变得非常容易

例

- 对于线性方程组求解问题：当系数矩阵A为下三角或者上三角时，方程组的求解变得非常容易
- A为一般矩阵时，可以首先将A分解为$PA = LU$，其中P为排列方阵，U和L分别是上、下三角阵

例

- 对于线性方程组求解问题：当系数矩阵A为下三角或者上三角时，方程组的求解变得非常容易
- A为一般矩阵时，可以首先将A分解为$PA = LU$，其中P为排列方阵，U和L分别是上、下三角阵
- 通过求解$Ly = Pb$，$Ux = y$就可以得到原方程的解

误差

误差来源于

- 原始数据的误差

误差

误差来源于

- 原始数据的误差
- 计算过程产生的误差

误差

误差来源于

- 原始数据的误差
- 计算过程产生的误差

- 误差是不可避免的

误差

误差来源于

- 原始数据的误差
- 计算过程产生的误差

- 误差是不可避免的
- 计算解与真解的差是多少？

敏度分解

研究原始数据的微小变化会引起解的多大变化。假设考虑函数$f(x)$的计算问题

- 中值定理与Taylor展开：可行性不大

敏度分解

数值代数基本问题

研究的必要性

主要技巧

误差分析

算法速度

考核方式

研究原始数据的微小变化会引起解的多大变化。假设考虑函数$f(x)$的计算问题

- 中值定理与Taylor展开：可行性不大
- 当$\delta x/|x|$很小时，确定一个尽可能小的正数$c(x)$，满足

$$\frac{|f(x+\delta x)-f(x)|}{|f(x)|} \leqslant c(x)\frac{|\delta x|}{|x|}$$

条件数

$c(x)$的大小在一定程度上反映了自变量的微小变化对函数值的影响程度

- $c(x)$称为f在x点的条件数

条件数

$c(x)$的大小在一定程度上反映了自变量的微小变化对函数值的影响程度

- $c(x)$称为f在x点的条件数
- $c(x)$很大时，自变量的小变化有可能引起函数值的大变化，因此称f在x点是病态的；
  当$c(x)$很小时，称f在x点是良态的

条件数

$c(x)$的大小在一定程度上反映了自变量的微小变化对函数值的影响程度

- $c(x)$称为f在x点的条件数
- $c(x)$很大时，自变量的小变化有可能引起函数值的大变化，因此称f在x点是病态的；
  当$c(x)$很小时，称f在x点是良态的
- 计算问题是否病态是问题自身的固有属性，与计算方法无关

舍入误差

- 计算机是有效精度的。这引起的误差称为舍入误差。

舍入误差

- 计算机是有效精度的。这引起的误差称为舍入误差。
- 分析舍入误差对算法结果的影响，是衡量算法优劣的重要标志

舍入误差

- 计算机是有效精度的。这引起的误差称为舍入误差。
- 分析舍入误差对算法结果的影响，是衡量算法优劣的重要标志
- 仍以函数求值为例

向后误差分析法

- 设采用某算法后函数f在x点的值为$\hat{y}$

向后误差分析法

- 设采用某算法后函数f在x点的值为$\hat{y}$
- 通过对每步具体运算的误差分析可以证明：存在δx满足$\hat{y} = f(x + \delta x)$, $|\delta x| \leqslant |x|\varepsilon$, 其中$\varepsilon$是一个与计算机精度和算法有关的正数

向后误差分析法

- 设采用某算法后函数f在x点的值为$\hat{y}$
- 通过对每步具体运算的误差分析可以证明：存在δx满足$\hat{y} = f(x + \delta x)$, $|\delta x| \leqslant |x|\varepsilon$, 其中$\varepsilon$是一个与计算机精度和算法有关的正数
- 这种把计算结果归结为原始数据经扰动后的精确结果的误差分析方法称为向后误差分析法

数值稳定性

- ε越小，说明舍入误差对算法的影响越小，因此称算法为数值稳定的，否则称为数值不稳定的

数值稳定性

- ε越小，说明舍入误差对算法的影响越小，因此称算法为数值稳定的，否则称为数值不稳定的
- 算法的数值稳定性是算法本身的固有属性，与计算问题是否病态无关

精度估计

数值代数基本问题

研究的必要性

主要技巧

误差分析

算法速度

考核方式

- 综合前述敏度分析和误差分析之后，计算结果的精度估计如下：

$$\frac{|\hat{y} - f(x)|}{|f(x)|} = \frac{|f(x+\delta x) - f(x)|}{|f(x)|} \leqslant c(x)\varepsilon$$

精度估计

- 综合前述敏度分析和误差分析之后，计算结果的精度估计如下：

$$\frac{|\hat{y} - f(x)|}{|f(x)|} = \frac{|f(x + \delta x) - f(x)|}{|f(x)|} \leqslant c(x)\varepsilon$$

- 计算结果是否可靠，依赖于计算问题是否病态以及所用的算法是否数值稳定

算法的种类

- 直接法：在没有误差的情况下可在有限步得到计算问题精确解

算法的种类

- 直接法：在没有误差的情况下可在有限步得到计算问题精确解
    - 运算量的大小可以作为其速度的一个主要标志

算法的种类

- 直接法：在没有误差的情况下可在有限步得到计算问题精确解
  - 运算量的大小可以作为其速度的一个主要标志
- 迭代法：采用逐次逼近的方法逼近问题的精确解，而在任意有限步都不能得到其精确解

算法的种类

- 直接法：在没有误差的情况下可在有限步得到计算问题精确解
  - 运算量的大小可以作为其速度的一个主要标志
- 迭代法：采用逐次逼近的方法逼近问题的精确解，而在任意有限步都不能得到其精确解
  - 除估计每步运算量的大小，还需要对收敛性进行分析

算法复杂性

计算或者估计算法的运算量

- 上世纪90年代之前：通常只计算乘除运算的次数

算法复杂性

计算或者估计算法的运算量

- 上世纪90年代之前：通常只计算乘除运算的次数
- 进入90年代之后：算法所有运算次数总和

算法复杂性

计算或者估计算法的运算量

- 上世纪90年代之前：通常只计算乘除运算的次数
- 进入90年代之后：算法所有运算次数总和
- 如果某一算法的运算量是关于n的多项式，通常就略去低阶项，用最高阶项称为算法的运算量

运算量与算法快慢

- 运算量只在一定程度上反映了算法的速度

运算量与算法快慢

- 运算量只在一定程度上反映了算法的速度
- 现代计算机中运算速度远远高于数据的传输速度，因此算法的速度在很大程度上依赖于算法实现后数据传输量的大小

收敛速度

针对迭代法

- 假设某一迭代法产生的序列$\{x_k\}$满足

$$\|x_k - x\| \leqslant c\|x_{k-1} - x\|^{\alpha}, \quad k = 1, 2, \dots$$

收敛速度

针对迭代法

- 假设某一迭代法产生的序列$\{x_k\}$满足

$$\|x_k - x\| \leqslant c\|x_{k-1} - x\|^{\alpha}, \quad k = 1, 2, \ldots$$

- 若$0 < c < 1$, $\alpha = 1$, 则称算法线性收敛

收敛速度

针对迭代法

- 假设某一迭代法产生的序列$\{x_k\}$满足

$$\|x_k - x\| \leqslant c\|x_{k-1} - x\|^{\alpha}, \quad k = 1, 2, \ldots$$

  - 若$0 < c < 1$, $\alpha = 1$, 则称算法线性收敛
  - 若$c > 0$, $\alpha = 2$, 则称算法平方收敛（二次收敛）。此时c越小越好

教材与参考书

- 教材：数值线性代数（第二版），徐树方等编著，北京大学出版社

教材与参考书

- 教材：数值线性代数（第二版），徐树方等编著，北京大学出版社
- 参考书

教材与参考书

- 教材：数值线性代数（第二版），徐树方等编著，北京大学出版社
- 参考书
    - 矩阵计算（第四版），G. H. Golub著，人民邮电出版社影印，2014年

教材与参考书

- 教材：数值线性代数（第二版），徐树方等编著，北京大学出版社
- 参考书
  - 矩阵计算（第四版），G. H. Golub著，人民邮电出版社影印，2014年
  - 代数特征值问题，J. H. Wilkinson著，科学出版社，2018年再版

考核方式

数值代数基本问题

研究的必要性

主要技巧

误差分析

算法速度

考核方式

- 课后习题：20% (注意网络上的解答有些是错的)

考核方式

数值代数基本问题
研究的必要性
主要技巧
误差分析
算法速度
考核方式

- 课后习题：20% (注意网络上的解答有些是错的)
- 编程作业：35% (课本上机习题+SVD分解)

考核方式

- 课后习题：20% (注意网络上的解答有些是错的)
- 编程作业：35% (课本上机习题+SVD分解)
- 期终考试：50% (闭卷?，课后习题和书上定理证明占60–70%)

考核方式

- 课后习题：20% (注意网络上的解答有些是错的)
- 编程作业：35% (课本上机习题+SVD分解)
- 期终考试：50% (闭卷？，课后习题和书上定理证明占60–70%)
- 所有作业在布置后的周三上交。特殊情况提前向助教说明，一周内补交有效！

考核方式

- 课后习题：20% (注意网络上的解答有些是错的)
- 编程作业：35% (课本上机习题+SVD分解)
- 期终考试：50% (闭卷?，课后习题和书上定理证明占60–70%)
- 所有作业在布置后的周三上交。特殊情况提前向助教说明，一周内补交有效!
- 发现照抄，双方均不得分!

考核方式

- 课后习题：20% (注意网络上的解答有些是错的)
- 编程作业：35% (课本上机习题+SVD分解)
- 期终考试：50% (闭卷？，课后习题和书上定理证明占60–70%)
- 所有作业在布置后的周三上交。特殊情况提前向助教说明，一周内补交有效！
- 发现照抄，双方均不得分！
- 以上各项可讨论，下周三确定！

编程作业的要求

- C++

编程作业的要求

- C++
- Python 3.6

编程作业的要求

- C++
- Python 3.6
- 具体编译平台，请在大群中与助教讨论确认。各语言最好只接受一种平台

编程作业的要求

- C++
- Python 3.6
- 具体编译平台，请在大群中与助教讨论确认。各语言最好只接受一种平台
- 课堂演示：C++