计算方法

Liquan Mei

School of Mathematics and Statistics, Xi'an Jiaotong University

数值计算方法

学习科学与工程计算中常用的数值计算方法及其相关理论。

学习科学与工程计算中常用的数值计算方法及其相关理论。

必要性 科学与技术的研究从定性前进到定量，科学计算与理论、实验三足鼎力，成为科学实践的三大手段。运用计算机进行技术是科技人员的基本技能，这要求研究掌握数值计算方法。

数值计算方法

学习科学与工程计算中常用的数值计算方法及其相关理论。

必要性 科学与技术的研究从定性前进到定量，科学计算与理论、实验三足鼎力，成为科学实践的三大手段。运用计算机进行技术是科技人员的基本技能，这要求研究掌握数值计算方法。

定量分析 工程问题 数学模型 选择计算方法 误差分析 异常解释

数值计算方法

学习科学与工程计算中常用的数值计算方法及其相关理论。

必要性 科学与技术的研究从定性前进到定量，科学计算与理论、实验三足鼎力，成为科学实践的三大手段。运用计算机进行技术是科技人员的基本技能，这要求研究掌握数值计算方法。

定量分析 工程问题　数学模型　选择计算方法　误差分析　异常解释

特点 数学抽象性　严密科学性　广泛应用性　高度技术性

是与计算机密切相连，实用性强的计算数学课程

学习科学与工程计算中常用的数值计算方法及其相关理论。

必要性 科学与技术的研究从定性前进到定量，科学计算与理论、实验三足鼎力，成为科学实践的三大手段。运用计算机进行技术是科技人员的基本技能，这要求研究掌握数值计算方法。

定量分析 工程问题　数学模型　选择计算方法　误差分析　异常解释

特点 数学抽象性　严密科学性　广泛应用性　高度技术性

是与计算机密切相连，实用性强的计算数学课程

任务 数值解　数值方法　数值分析（可靠性、效率）

学习科学与工程计算中常用的数值计算方法及其相关理论。

必要性 科学与技术的研究从定性前进到定量，科学计算与理论、实验三足鼎力，成为科学实践的三大手段。运用计算机进行技术是科技人员的基本技能，这要求研究掌握数值计算方法。

定量分析 工程问题　数学模型　选择计算方法　误差分析　异常解释

特点 数学抽象性　严密科学性　广泛应用性　高度技术性

是与计算机密切相连，实用性强的计算数学课程

任务 数值解　数值方法　数值分析（可靠性、效率）

重点 方法的构造和使用　工作量　收敛性　优缺点

数值计算问题的类型

1. 离散问题如求解方程组
2. 连续问题的离散化如数值积分、数值微分、常微分方程数值解、偏微分方程数值解
3. 离散问题的连续化数值拟合、数据逼近

课程基础 高等数学, 线性代数,计算机语言,数据结构

参考书 1.李乃成,梅立泉, 数值分析, 科学出版社;

2.Alfio Quarteroni, Riccardo Sacco, Fausto Saleri. Numerical Mathematics. Springer Science Business Media, Inc.2000.

3.邓建中, 刘之行. 计算方法. 西安:西安交通大学出版社. 2001

成绩 考试成绩80%,上机作业20%

网页

http://cc.xjtu.edu.cn/G2S/site/preview#/home/v?currentoc=363

http://gr.xjtu.edu.cn/web/lqmei

要求

1. 了解数值计算方法一般**概念**
2. 了解**误差的来源与分类**、数据误差的影响
3. 舍入误差对数值计算的影响
4. **计算中应注意的原则**

要求

1. 了解数值计算方法一般**概念**
2. 了解**误差的来源与分类**、数据误差的影响
3. 舍入误差对数值计算的影响
4. **计算中应注意的原则**

一般概念

算法 由基本运算及运算顺序的规定构成的解题步骤

要求

1 了解数值计算方法一般**概念**

2 了解**误差的来源与分类**、数据误差的影响

3 舍入误差对数值计算的影响

4 **计算中应注意的原则**

一般概念

算法 由基本运算及运算顺序的规定构成的解题步骤

好算法 运算量少 存储小 逻辑结构简单 易编程 结果可靠

要求

1 了解数值计算方法一般**概念**

2 了解**误差的来源与分类**、数据误差的影响

3 舍入误差对数值计算的影响

4 **计算中应注意的原则**

一般概念

算法 由基本运算及运算顺序的规定构成的解题步骤

好算法 运算量少 存储小 逻辑结构简单 易编程 结果可靠

稳定性 计算过程中误差可控，各步误差对结果不致产生过大影响

要求

1 了解数值计算方法一般**概念**

2 了解**误差的来源与分类**、数据误差的影响

3 舍入误差对数值计算的影响

4 **计算中应注意的原则**

一般概念

算法 由基本运算及运算顺序的规定构成的解题步骤

好算法 运算量少 存储小 逻辑结构简单 易编程 结果可靠

稳定性 计算过程中误差可控，各步误差对结果不致产生过大影响

收敛性 增加计算量，近似解充分接近真解

误差

来源与分类

1. 模型误差: 在建立数学模型时，忽略次要因素而造成的
2. 观测误差、数据误差或参量误差:
3. 截断误差或方法误差
4. 舍入误差或计算误差:由于计算机中的性能限制而造成的

误差

来源与分类

1. 模型误差: 在建立数学模型时，忽略次要因素而造成的
2. 观测误差、数据误差或参量误差:
3. 截断误差或方法误差
4. 舍入误差或计算误差:由于计算机中的性能限制而造成的

设x为准确值, $\tilde{x}$是x 的一个近似值,

- **绝对误差** $\Delta x = x - \tilde{x}$ 或 $|\Delta x| = |x - \tilde{x}| \le \epsilon \quad x = \tilde{x} \pm \epsilon$

误差

来源与分类

1. 模型误差: 在建立数学模型时，忽略次要因素而造成的
2. 观测误差、数据误差或参量误差:
3. 截断误差或方法误差
4. 舍入误差或计算误差:由于计算机中的性能限制而造成的

设x为准确值, $\tilde{x}$是x 的一个近似值,

- **绝对误差** $\Delta x = x - \tilde{x}$ 或 $|\Delta x| = |x - \tilde{x}| \leq \epsilon \quad x = \tilde{x} \pm \epsilon$
- **相对误差** $\delta x = \frac{x-\tilde{x}}{x}$ 或 $|\delta x|$

[◇]**准确数字** $|x - \tilde{x}| \leq \frac{1}{2} \times 10^{-n}$, 则称近似值$\tilde{x}$ 准确到 n 位小数，并将从该位起直到最左端的非零数字之间的一切数字称为近似值$\tilde{x}$ 的**准确数字（有效数字）**

[◇]**准确数字** $|x-\tilde{x}| \leq \frac{1}{2}\times 10^{-n}$, 则称近似值$\tilde{x}$ 准确到 n 位小数，并将从该位起直到最左端的非零数字之间的一切数字称为近似值$\tilde{x}$ 的准确数字（有效数字）

如π，若取π 的近似值分别为$\pi_1 = 3.1416$, $\pi_2 = \frac{22}{7} = 3.1428\ldots$

[◇]**准确数字** $|x-\tilde{x}| \le \frac{1}{2}\times 10^{-n}$, 则称近似值$\tilde{x}$ 准确到 n 位小数，并将从该位起直到最左端的非零数字之间的一切数字称为近似值$\tilde{x}$ 的准确数字（有效数字）

如π，若取π 的近似值分别为$\pi_1 = 3.1416$, $\pi_2 = \frac{22}{7} = 3.1428\ldots$.则

$|\pi-\pi_1| = 0.000\,007\,3\cdots < 0.000\,05 = \frac{1}{2}\times 10^{-4}$,

$|\pi-\pi_2| = 0.0012\cdots < 0.005 = \frac{1}{2}\times 10^{-2}$.

由此可知近似值π_1 准确到4位小数，具有5位准确数字.

[$\diamond$]**准确数字** $|x-\tilde{x}| \leq \frac{1}{2} \times 10^{-n}$，则称近似值$\tilde{x}$ 准确到 n 位小数，并将从该位起直到最左端的非零数字之间的一切数字称为近似值$\tilde{x}$ 的**准确数字（有效数字）**

[$\diamond$]**准确数字** $|x-\tilde{x}| \le \frac{1}{2}\times 10^{-n}$, 则称近似值$\tilde{x}$ 准确到 n 位小数，并将从该位起直到最左端的非零数字之间的一切数字称为近似值$\tilde{x}$ 的**准确数字（有效数字）**

若把$x \quad (x_1 \neq 0)$写成标浮点形式

$$\tilde{x} = \pm x_1 x_2 \cdots x_m \,.\, x_{m+1} \cdots x_{m+n} = \pm 10^m \times 0\,.\,x_1 x_2 \cdots x_{n+m}$$

绝对误差界为$\frac{1}{2}\times 10^{-n}$ 时，$\tilde{x}$准确到n 位小数，具有$n+m$ 位准确数字.

$$\tilde{x} = \pm 0.0 \cdots 0 x_1 x_2 \cdots x_{n-m} = \pm 10^{-m} \times 0\,.\,x_1 x_2 \cdots x_{n-m}$$

绝对误差界为$\frac{1}{2}\times 10^{-n}$ 时，$\tilde{x}$准确到n 位小数，具有$n-m$ 位准确数字

1-1 已知$\sqrt{2} = 1.414\,213\,562\,373\cdots$，分别写出准确到3至5位小数的近似值，指出它们的绝对误差界，相对误差界以及有效数字的位数.

解：

$$\sqrt{2} \approx \tilde{x}_1 = 1.414, \qquad \sqrt{2} \approx \tilde{x}_2 = 1.414\,2,$$
$$|\sqrt{2} - 1.414| \le \frac{1}{2} \times 10^{-3}, \quad |\sqrt{2} - 1.414\,2| \le \tfrac{1}{2} \times 10^{-4},$$

$$\delta_1 = \frac{|\sqrt{2} - 1.414|}{\sqrt{2}} \le \frac{\frac{1}{2} \times 10^{-3}}{\sqrt{2}} \approx 0.353\,553\,39 \times 10^{-3}$$
$$\delta_2 = \frac{|\sqrt{2} - 1.414\,2|}{\sqrt{2}} \le \frac{\frac{1}{2} \times 10^{-4}}{\sqrt{2}} \approx 0.353\,553\,39 \times 10^{-4}$$

计算机中的实数采用浮点表示法，即将一个数分为指数和尾数两部分来表示. 设计算机采用 β 进制

$$fl(x) = \tilde{x} = \pm\{\frac{x_1}{\beta} + \frac{x_2}{\beta^2} + \cdots + \frac{x_t}{\beta^t}\} \times \beta^l = \pm 0\,.\,x_1 x_2 \cdots x_t \times \beta^l,$$

其中$x_1 \in \{1, 2, \cdots, \beta - 1\}$，$x_i \in \{0, 1, \cdots, \beta - 1\}$，$i = 2, 3, \cdots, t$.
指数l 是整数，也称为阶码. 阶码l 的取值范围
为$L \leq l \leq U (L < 0, U > 0)$.

单精度浮点数按32位存储，双精度浮点数按64位存储。单精度浮点数的尾数为23位、阶数为8位；双精度浮点数的尾数为53位（包含符号位）、阶数为11位（包含符号位）。

计算机所能表示的全部浮点数的集合称为计算机的**浮点数集**，记为

$$F(\beta, t, L, U) = \{0\} \bigcup \left\{ fl(x) = \pm 0\,.\,x_1 x_2 \cdots x_t \times \beta^l \right\}.$$

浮点数集中共有$2(\beta-1)\beta^{t-1}(U-L+1)+1$ 个数，它所能表示的数的范围为

$$fl_{\min}(x) = \beta^{L-1} \leq \mid fl(x) \mid \leq \beta^U \times (1-\beta^{-t}) = fl_{\max}(x).$$

浮点数的特点

- 实数转换到浮点数–浮点化，〈缺点：〉会产生误差, 按四舍五入，绝对误差. 〈优点：〉浮点化产生的相对误差有界一个实数x一般只能用最接近的浮点数$fl(x)$替代它,由此将产生舍入误差. 浮点数$fl(x)$ 的绝对误差与相对误差分别为

$$|x - fl(x)| \leq \frac{1}{2}\beta^{-t} \times \beta^{l} = \frac{1}{2}\beta^{l-t} \ ,$$

$$\frac{|x - fl(x)|}{|\ x\ |} \leq \frac{\frac{1}{2}\beta^{l-t}}{\beta^{l-1}} = \frac{1}{2}\beta^{-(t-1)} \ .$$

  相对误差限$\frac{1}{2}\beta^{-(t-1)}$ 只与计算机数的进制和字长有关，称为**计算机的相对精度**.
- 每一个浮点数集的数字有限
- 浮点数集中的运算非自封闭,（因为数字有限等）

浮点数

- 浮点数运算结果的指数I不在范围[L,U]中，例如：F(2,3,-1,2)中，$(0.100 * 2^2) \times (0.110 * 2^2) = 0.110 * 2^3$. 上溢
- 结果的尾数多于t位数字，例如：F(2,3,-1,2)中，$(0.100 * 2^0) + (0.111 * 2^0) = 0.1011 * 2^1$ 舍入.
- 在浮点数集中数据的尾数字长t是有限,当两个相近数相减时,会损失比较多的有效数字
- 在相同的指数条件下，两个数量相差较大的数字相加（减）时，较小数的有效数字会被丧失

计算中应注意的原则

1. **尽量避免两个相近的数相减. 常用的做法是变换算式.**
例如，当$|x|$的绝对值充分大时，$\sqrt{x+1}-\sqrt{x}=\frac{1}{\sqrt{x+1}+\sqrt{x}}$.

计算中应注意的原则

1. **尽量避免两个相近的数相减．常用的做法是变换算式．**
   例如，当$|x|$的绝对值充分大时，$\sqrt{x+1}-\sqrt{x}=\frac{1}{\sqrt{x+1}+\sqrt{x}}$.
2. **尽量避免绝对值很大的数作乘数,绝对值较小的数作除数.**
   由绝对误差估计式可知两数乘、除运算的绝对误差分别为：

$$\Delta(x_1x_2)\approx x_2\Delta x_1+x_1\Delta x_2\ ,\Delta(\frac{x_1}{x_2})\approx\frac{\Delta x_1}{x_2}-\frac{x_1}{x_2^2}\Delta x_2\ .$$

计算中应注意的原则

1. **尽量避免两个相近的数相减．常用的做法是变换算式.**
   例如，当$|x|$的绝对值充分大时，$\sqrt{x+1}-\sqrt{x}=\frac{1}{\sqrt{x+1}+\sqrt{x}}$.
2. **尽量避免绝对值很大的数作乘数,绝对值较小的数作除数.**
   由绝对误差估计式可知两数乘、除运算的绝对误差分别为:

$$\Delta(x_1x_2)\approx x_2\Delta x_1+x_1\Delta x_2\ ,\Delta(\frac{x_1}{x_2})\approx\frac{\Delta x_1}{x_2}-\frac{x_1}{x_2^2}\Delta x_2\ .$$

3. **若干个数相加,应采取绝对值较小者先加的原则.**

计算中应注意的原则

1 **尽量避免两个相近的数相减. 常用的做法是变换算式.**
例如，当$|x|$的绝对值充分大时，$\sqrt{x+1}-\sqrt{x}=\frac{1}{\sqrt{x+1}+\sqrt{x}}$.

2 **尽量避免绝对值很大的数作乘数,绝对值较小的数作除数.**
由绝对误差估计式可知两数乘、除运算的绝对误差分别为：

$$\Delta(x_1x_2)\approx x_2\Delta x_1+x_1\Delta x_2\ ,\Delta(\frac{x_1}{x_2})\approx\frac{\Delta x_1}{x_2}-\frac{x_1}{x_2^2}\Delta x_2\ .$$

3 **若干个数相加,应采取绝对值较小者先加的原则.**

4 **尽量减少乘除法的运算次数.**
例如计算多项式$p_n(x)=a_0+a_1x+a_2x^2+\cdots+a_nx^n$ 的值，若先计算各项a_kx^k ，再逐项相加，则需进行$1+2+3\cdots+n=\frac{n(n+1)}{2}$ 次乘法和n次加法. 若用秦九韶算法，将$p_n(x)$改写为

$$p_n(x)=(\cdots((a_nx+a_{n-1})x+a_{n-2})x+\cdots+a_1)x+a_0\ .$$

则只需n 次乘法和n 次加法.

问题的性态

- 所计算的问题当原始数据发生小扰动时，问题的解一般也发生扰动。
- 问题的性态–问题的解对原始数据发生变化的敏感性。
- 原始数据小扰动导致问题解小扰动，称为**良态问题**；
- 反之，原始数据小扰动导致问题解大扰动，称为**病态问题**

问题的性态

- 所计算的问题当原始数据发生小扰动时，问题的解一般也发生扰动。
- 问题的性态–问题的解对原始数据发生变化的敏感性。
- 原始数据小扰动导致问题解小扰动，称为**良态问题**；
- 反之，原始数据小扰动导致问题解大扰动，称为**病态问题**

条件数 设原始数据x, 计算结果$f(x)$ ，扰动后的数据$\tilde{x}$ 计算结果$f(\tilde{x})$，若问题f 存在常数m，满足关系式

$$|\frac{f(x)-f(\tilde{x})}{f(x)}| \leq m|\frac{x-\tilde{x}}{x}|$$

或则称（相对误差之比的上界）m为该问题的条件数，记作$cond(f)$.

例如，计算积分

$$I_k = e^{-1}\int_0^1 x^k e^x dx, \quad k = 0, 1, \cdots, 7 \ .$$

例如，计算积分

$$I_k = e^{-1}\int_0^1 x^k e^x dx, \quad k = 0, 1, \cdots, 7 \ .$$

算法1 $I_0 = e^{-1}\int_0^1 e^x dx = 1 - \frac{1}{e}$,,

$$I_k = 1 - kI_{k-1}, \quad k = 1, 2, \cdots, 7 \ .$$

例如，计算积分

$$I_k = e^{-1}\int_0^1 x^k e^x dx, \quad k = 0, 1, \cdots, 7 \ .$$

算法1 $I_0 = e^{-1}\int_0^1 e^x dx = 1 - \frac{1}{e}$,,

$$I_k = 1 - kI_{k-1}, \quad k = 1, 2, \cdots, 7 \ .$$

算法2 $I_k = \int_0^1 x^k e^{-(1-x)} dx < \int_0^1 x^k dx = \frac{1}{k+1} \to 0$, $I_{k-1} = (1 - I_k)/k$, 取 $I_{11} = 0$.

例如，计算积分

$$I_k = e^{-1}\int_0^1 x^k e^x dx, \quad k = 0, 1, \cdots, 7 \ .$$

算法1 $I_0 = e^{-1}\int_0^1 e^x dx = 1 - \frac{1}{e}$ „

$$I_k = 1 - kI_{k-1}, \quad k = 1, 2, \cdots, 7 \ .$$

算法2 $I_k = \int_0^1 x^k e^{-(1-x)} dx < \int_0^1 x^k dx = \frac{1}{k+1} \to 0,$
$I_{k-1} = (1 - I_k)/k$, 取$I_{11} = 0$.

I_k	I_0	I_1	I_2	I_3	I_5	I_7
准确值	0.632 1	0.367 9	0.264 2	0.207 3	0.145 5	0.112 4
方法 1	0.632 1	0.367 9	0.264 2	0.207 4	0.148 0	0.216 0
方法 2	0.632 1	0.367 9	0.264 2	0.207 3	0.145 5	0.112 4

方法①中，原始步的误差，随着计算步数的增加被严重地放大

要求

1. 熟练掌握高斯消去法及选列主元技术
2. 掌握三角分解法
3. 掌握向量范数与矩阵范数、方程组的条件数，了解舍入误差对解的影响

要求

1. 熟练掌握高斯消去法及选列主元技术
2. 掌握三角分解法
3. 掌握向量范数与矩阵范数、方程组的条件数，了解舍入误差对解的影响

$Ax = b$, 系数矩阵的行列式$|A| \neq 0$，则方程组的解存在且唯一.

直接法:

要求

1. 熟练掌握高斯消去法及选列主元技术
2. 掌握三角分解法
3. 掌握向量范数与矩阵范数、方程组的条件数，了解舍入误差对解的影响

$Ax = b$，系数矩阵的行列式$|A| \neq 0$，则方程组的解存在且唯一.
直接法： 在无舍入误差的假定下，经过有限次运算就可以求得方程组的准确解.
基本原理：

要求

1. 熟练掌握高斯消去法及选列主元技术
2. 掌握三角分解法
3. 掌握向量范数与矩阵范数、方程组的条件数，了解舍入误差对解的影响

$Ax = b$, 系数矩阵的行列式$|A| \neq 0$，则方程组的解存在且唯一.
直接法： 在无舍入误差的假定下，经过有限次运算就可以求得方程组的准确解.
基本原理： $Ax = b$（一般）　**等价变换**　$Gx = d$ (结构简单)
G: 对角阵、三角阵、结构简单阵

高斯消去法

高斯消去法首先是将方程组进行消元运算，将其化为一个等价的同解的上三角方程组，这个过程称为消元过程. 然后通过求解上三角方程组得到原方程组的解，后一过程称为回代过程.

高斯消去法首先是将方程组进行消元运算，将其化为一个等价的同解的上三角方程组，这个过程称为消元过程. 然后通过求解上三角方程组得到原方程组的解，后一过程称为回代过程.

例2.1 用高斯消去法求解线性方程组

$$\begin{cases} x_1 + 2x_2 + x_3 = 0, \\ 2x_1 + 2x_2 + 3x_3 = 3, \\ -x_1 - 3x_2 \qquad = 2. \end{cases}$$

高斯消去法首先是将方程组进行消元运算，将其化为一个等价的同解的上三角方程组，这个过程称为消元过程. 然后通过求解上三角方程组得到原方程组的解，后一过程称为回代过程.

例2.1 用高斯消去法求解线性方程组

$$\begin{cases} x_1 + 2x_2 + x_3 = 0, \\ 2x_1 + 2x_2 + 3x_3 = 3, \\ -x_1 - 3x_2 \quad = 2. \end{cases}$$

解 消元过程：

第1步 保留第1个方程不动，将第2、3个方程中x_1前的系数消为零.为此，第2、3个方程分别减去第1个方程的2、-1倍，得

$$\begin{aligned} x_1 + 2x_2 + x_3 &= 0, \\ -2x_2 + x_3 &= 3, \\ -x_2 + x_3 &= 2. \end{aligned}$$

第2步　保留第1、2个方程不动，将第3个方程中x_2前的系数消为零. 第3个方程减去第2个方程的$\frac{1}{2}$倍，则得与原方程同解的上三角方程组：

$$\begin{aligned} x_1 + 2x_2 + x_3 &= 0, \\ -2x_2 + x_3 &= 3, \\ \frac{1}{2}x_3 &= \frac{1}{2}. \end{aligned}$$

第2步　保留第1、2个方程不动，将第3个方程中x_2前的系数消为零. 第3个方程减去第2个方程的$\frac{1}{2}$倍，则得与原方程同解的上三角方程组：

$$\begin{aligned} x_1 + 2x_2 + x_3 &= 0, \\ -2x_2 + x_3 &= 3, \\ \frac{1}{2}x_3 &= \frac{1}{2}. \end{aligned}$$

回代过程:
求解以上上三角方程组. 从第3个方程组求得$x_3 = 1$，代入第2个方程可得$x_2 = -1$，再将x_2, x_3代入第1个方程组得$x_1 = 1$.

为了便于高斯消去法在计算机上的实现以及减少计算机的存贮量，在开始计算时将方程组的系数矩阵存放在二维数组$A=(a_{ij})_{n\times n}$中，方程的右端项存放在一维数组b中．消去过程中，记$l_{ik}=a_{ik}^{(k-1)}/a_{kk}^{(k-1)}$，$a_{ij}^{(k)}=a_{ij}^{(k-1)}-l_{ik}a_{kj}^{(k-1)}$，$(i=k+1,k+2,\cdots,n,\quad j=k+1,k+2,\cdots,n)$．消元过程结束后，二维数组$A$和一维数组$b$的内容如下：

$$A=\begin{pmatrix} a_{11}^{(0)} & a_{12}^{(0)} & a_{13}^{(0)} & \cdots & a_{1,n-1}^{(0)} & a_{1n}^{(0)} \\ l_{21} & a_{22}^{(1)} & a_{23}^{(1)} & \cdots & a_{2,n-1}^{(1)} & a_{2n}^{(1)} \\ l_{31} & l_{32} & a_{33}^{(2)} & \cdots & a_{3,n-1}^{(2)} & a_{3n}^{(2)} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ l_{n1} & l_{n2} & l_{n3} & \cdots & l_{n,n-1} & a_{nn}^{(n-1)} \end{pmatrix},\ b=\begin{pmatrix} b_1^{(0)} \\ b_2^{(1)} \\ b_3^{(2)} \\ \cdots \\ b_n^{(n-1)} \end{pmatrix}.$$

高斯消去法的算法组织

高斯消去法的计算过程

消去

$for \quad k = 1 \ to \ n-1 \ do$

$\quad if \quad |a_{kk}| \le \epsilon, \ then \quad stop$

$\quad for \quad i = k+1 \ to \ n \ do$

$\quad\quad a_{ik} := a_{ik}/a_{kk}$

$\quad\quad for \quad j = k+1 \ to \ n \ do$

$\quad\quad\quad a_{ij} := a_{ij} - a_{ik}a_{kj}$

$\quad\quad end \ do$

$\quad\quad b_i := b_i - a_{ik}b_k$

$\quad end \ do$

$end \ do$

高斯消去法的算法组织

高斯消去法的计算过程

消去

$$
\begin{array}{l}
\textit{for} \quad k = 1 \ \textit{to} \ n-1 \ \textit{do} \\
\quad \textit{if} \quad |a_{kk}| \le \epsilon, \ \textit{then} \quad \textit{stop} \\
\quad \textit{for} \quad i = k+1 \ \textit{to} \ n \ \textit{do} \\
\quad\quad a_{ik} := a_{ik}/a_{kk} \\
\quad\quad \textit{for} \quad j = k+1 \ \textit{to} \ n \ \textit{do} \\
\quad\quad\quad a_{ij} := a_{ij} - a_{ik}a_{kj} \\
\quad\quad \textit{end do} \\
\quad\quad b_i := b_i - a_{ik}b_k \\
\quad \textit{end do} \\
\textit{end do}
\end{array}
$$

回代过程

$$
\begin{array}{l}
x_n := b_n/a_{nn} \\
\textit{for} \quad k = n-1 \ \textit{to} \ 1 \quad -1 \ \textit{do} \\
\quad s := b_k \\
\quad \textit{for} \quad j = k+1 \ \textit{to} \ n \ \textit{do} \\
\quad\quad s := s - a_{kj}x_j \\
\quad \textit{end do} \\
\quad x_k := s/a_{kk} \\
\textit{end do}
\end{array}
$$

高斯消去法中乘除法的运算量

消元过程共进行$n-1$步，即$k=1,2,\cdots,n-1$.
则消元过程中乘除法的运算量为

$$N_1=\sum_{k=1}^{n-1}[(n-k)^2+2(n-k)]=\frac{n^3}{3}+\frac{n^2}{2}-\frac{5n}{6}.$$

回代过程乘除法的运算量为

$$N_2=\sum_{k=1}^{n-1}(n-k)+n=\frac{n^2}{2}+\frac{n}{2}.$$

所以，用高斯消去法解n阶线性方程组的乘除法运算量为

$$N=N_1+N_2=\frac{n^3}{3}+\frac{n^2}{2}-\frac{5}{6}n+\frac{n^2}{2}+\frac{n}{2}=\frac{1}{3}n^3+n^2-\frac{1}{3}n.$$

高斯消去法顺利进行的条件

从消元过程知，当且仅当$a_{kk}^{(k-1)} \neq 0 \quad (k = 1, 2, \cdots, n)$时，高斯消去法才能进行下去.

高斯消去法能顺利进行的几个条件分别是：

高斯消去法顺利进行的条件

从消元过程知，当且仅当$a_{kk}^{(k-1)} \neq 0 \quad (k = 1, 2, \cdots, n)$时，高斯消去法才能进行下去.

高斯消去法能顺利进行的几个条件分别是:

条件1 系数矩阵A的各阶顺序主子式均不等于零.

$$D_1 = a_{11} \neq 0, \quad D_k = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{vmatrix} \neq 0, \quad k = 2, \cdots, n.$$

高斯消去法顺利进行的条件

从消元过程知，当且仅当$a_{kk}^{(k-1)} \neq 0 \quad (k=1,2,\cdots,n)$时，高斯消去法才能进行下去.

高斯消去法能顺利进行的几个条件分别是：

条件1 系数矩阵A的各阶顺序主子式均不等于零.

$$D_1 = a_{11} \neq 0, \quad D_k = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{vmatrix} \neq 0, \quad k=2,\cdots,n.$$

条件2 系数矩阵A是对称正定矩阵.

从消元过程知，当且仅当$a_{kk}^{(k-1)} \neq 0 \quad (k=1,2,\cdots,n)$时，高斯消去法才能进行下去.

高斯消去法能顺利进行的几个条件分别是:

条件1 系数矩阵A的各阶顺序主子式均不等于零.

$$D_1 = a_{11} \neq 0, \quad D_k = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{vmatrix} \neq 0, \quad k = 2,\cdots,n.$$

条件2 系数矩阵A是对称正定矩阵.

条件3 系数矩阵A是严格对角占优阵.

$$\sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}| < |a_{ii}|, \quad i = 1,2,\cdots,n.$$

在高斯消去法的消元过程中要求主元$a_{kk}^{(k-1)} \neq 0$.
列主元高斯消去法:
在第k步消元之前，从$A^{(k-1)}$的第k列元素$a_{kk}^{(k-1)}, \cdots, a_{nk}^{(k-1)}$ 中选取绝对值最大的元$a_{pk}^{(k-1)}$，交换交换第p 个方程与第k个方程的位置，使新的$a_{kk}^{(k-1)}$的绝对值最大，然后进行第k步消元.

$$
\text{选主元素} \quad \left\{
\begin{array}{l}
max := 0.0 \\
for \quad i = k \ \ to \ \ n \ \ do \\
\quad if \quad |a_{ik}| \leq |max| \quad then \\
\quad\quad go \ \ to \quad \text{标号}10 \\
\quad else \quad max := a_{ik} \\
\quad q := i \\
\text{标号}10 \quad end \ \ do
\end{array}
\right.
$$

在高斯消去法的消元过程中要求主元$a_{kk}^{(k-1)} \neq 0$.
列主元高斯消去法:
在第k步消元之前，从$A^{(k-1)}$的第k列元素$a_{kk}^{(k-1)}, \cdots, a_{nk}^{(k-1)}$中选取绝对值最大的元$a_{pk}^{(k-1)}$，交换交换第$p$个方程与第$k$个方程的位置，使新的$a_{kk}^{(k-1)}$的绝对值最大，然后进行第$k$步消元.

$$
\text{选主元素} \quad \left\{ \begin{array}{l}
max := 0.0 \\
for \quad i = k \ \ to \ \ n \ \ do \\
\quad if \quad |a_{ik}| \leq |max| \quad then \\
\qquad go \ \ to \quad \text{标号}10 \\
\quad else \quad max := a_{ik} \\
\quad q := i \\
\text{标号}10 \quad end \ \ do
\end{array} \right.
$$

$$
\begin{array}{l}
if \quad |max| \leq \epsilon, \ \ then \quad stop \\
if \quad q = k \quad then \quad go \ \ to \quad \text{消元部分}
\end{array}
$$

列主元高斯消去法

在高斯消去法的消元过程中要求主元$a_{kk}^{(k-1)} \neq 0$.

列主元高斯消去法:

在第k步消元之前，从$A^{(k-1)}$的第k列元素$a_{kk}^{(k-1)}, \cdots, a_{nk}^{(k-1)}$ 中选取绝对值最大的元$a_{pk}^{(k-1)}$，交换交换第p 个方程与第k个方程的位置，使新的$a_{kk}^{(k-1)}$的绝对值最大，然后进行第k步消元.

选主元素
$$\left\{\begin{array}{l} max := 0.0 \\ for \quad i = k \ to \ n \ do \\ \quad if \quad |a_{ik}| \leq |max| \quad then \\ \quad\quad go \ to \quad \text{标号}10 \\ \quad else \quad max := a_{ik} \\ \quad q := i \\ \text{标号}10 \quad end \ do \end{array}\right.$$

交换方程位置
$$\left\{\begin{array}{l} for \quad j = k \ to \ n \ do \\ \quad c := a_{kj} \\ \quad a_{kj} := a_{qj} \\ \quad a_{qj} := c \\ end \ do \\ c := b_k \\ b_k := b_q \\ b_q := c \end{array}\right.$$

$if \quad |max| \leq \epsilon, \ then \quad stop$

$if \quad q = k \quad then \quad go \ to$ 消元部分

例　用列主元素高斯消去法解方程组

$$\begin{cases} x_1 + 2x_2 + x_3 = 0, \\ 2x_1 + 2x_2 + 3x_3 = 3, \\ -x_1 - 3x_2 \quad = 2. \end{cases}$$

例 用列主元素高斯消去法解方程组

$$\begin{cases} x_1 + 2x_2 + x_3 = 0, \\ 2x_1 + 2x_2 + 3x_3 = 3, \\ -x_1 - 3x_2 \quad = 2. \end{cases}$$

解 消元过程

第1步

在系数矩阵A的第1列元素中选取绝对值最大的元$a_{21} = 2$,交换第1个方程与第2个方程的位置，得

$$\begin{cases} 2x_1 + 2x_2 + 3x_3 = 3, & (1) \\ x_1 + 2x_2 + x_3 = 0, & (2) \\ -x_1 - 3x_2 \quad = 2. & (3) \end{cases}$$

例 用列主元素高斯消去法解方程组

$$\begin{cases} x_1 + 2x_2 + x_3 = 0, \\ 2x_1 + 2x_2 + 3x_3 = 3, \\ -x_1 - 3x_2 \quad = 2. \end{cases}$$

解 消元过程

第1步

在系数矩阵A的第1列元素中选取绝对值最大的元$a_{21} = 2$,交换第1个方程与第2个方程的位置，得

$$\begin{cases} 2x_1 + 2x_2 + 3x_3 = 3, & (1) \\ x_1 + 2x_2 + x_3 = 0, & (2) \\ -x_1 - 3x_2 \quad = 2. & (3) \end{cases}$$

进行第1步消元，方程(2)−方程(1) $\times \frac{1}{2}$, 方程(3)−方程(1) $\times \left(-\frac{1}{2}\right)$, 得

$$\begin{cases} 2x_1 + 2x_2 + 3x_3 = 3, & (1') \\ x_2 - \frac{1}{2}x_3 = -\frac{3}{2}, & (2') \\ -2x_2 + \frac{3}{2}x_3 = \frac{7}{2}. & (3') \end{cases}$$

第2步

在第1步消元后的矩阵$A^{(1)}$的元素$a_{22}^{(1)}=1$，$a_{32}^{(1)}=-2$ 中选取绝对值最大的元$a_{32}^{(1)}=-2$作为主元素，交换第2个方程与第3个方程的位置，得

$$\begin{cases} 2x_1+2x_2+3x_3=3, & (1'') \\ \quad -2x_2+\frac{3}{2}x_3=\frac{7}{2}, & (2'') \\ \quad x_2-\frac{1}{2}x_3=-\frac{3}{2}. & (3'') \end{cases}$$

第2步

在第1步消元后的矩阵$A^{(1)}$的元素$a_{22}^{(1)}=1$，$a_{32}^{(1)}=-2$ 中选取绝对值最大的元$a_{32}^{(1)}=-2$作为主元素，交换第2个方程与第3个方程的位置，得

$$\begin{cases} 2x_1+2x_2+3x_3=3, & (1'') \\ -2x_2+\frac{3}{2}x_3=\frac{7}{2}, & (2'') \\ x_2-\frac{1}{2}x_3=-\frac{3}{2}. & (3'') \end{cases}$$

进行第2步消元，

方程$(3'')-$方程$(2'')\times\left(-\frac{1}{2}\right)$，得

$$\begin{cases} 2x_1+2x_2+3x_3=3, \\ -2x_2+\frac{3}{2}x_3=\frac{7}{2}, \\ \frac{1}{4}x_3=\frac{1}{4}. \end{cases}$$

第2步

在第1步消元后的矩阵$A^{(1)}$的元素$a_{22}^{(1)}=1$, $a_{32}^{(1)}=-2$ 中选取绝对值最大的元$a_{32}^{(1)}=-2$作为主元素，交换第2个方程与第3个方程的位置，得

$$\begin{cases} 2x_1+2x_2+3x_3=3, & (1'') \\ -2x_2+\frac{3}{2}x_3=\frac{7}{2}, & (2'') \\ x_2-\frac{1}{2}x_3=-\frac{3}{2}. & (3'') \end{cases}$$

进行第2步消元，

方程$(3'')$–方程$(2'')\times(-\frac{1}{2})$，得

$$\begin{cases} 2x_1+2x_2+3x_3=3, \\ -2x_2+\frac{3}{2}x_3=\frac{7}{2}, \\ \frac{1}{4}x_3=\frac{1}{4}. \end{cases}$$

由回代过程求得原方程组的解$x=(1,-1,1)^T$.

矩阵的三角分解

在实际问题中，经常遇到系数矩阵A相同，但右端项不同的多个线性方程组$Ax = b^{(i)}$ $(i = 1, 2, \cdots, m)$. **矩阵的三角分解法**.

矩阵的三角分解

在实际问题中，经常遇到系数矩阵A相同，但右端项不同的多个线性方程组$Ax = b^{(i)}$ $(i = 1, 2, \cdots, m)$. **矩阵的三角分解法**.

从矩阵运算的观点看高斯消去法的消元过程实质上是将增广矩阵$(A^{(0)}, b^{(0)})$通过逐步左乘一系列的初等下三角矩阵$L_k(k = 1, \cdots, n-1)$，最终变为矩阵$(A^{(n-1)}, b^{(n-1)})$, 即将原方程组化为一个同解的上三角方程组$A^{(n-1)}x = b^{(n-1)}$.

在实际问题中，经常遇到系数矩阵A相同，但右端项不同的多个线性方程组$Ax = b^{(i)}$ $(i = 1, 2, \cdots, m)$. **矩阵的三角分解法**.

从矩阵运算的观点看高斯消去法的消元过程实质上是将增广矩阵$(A^{(0)}, b^{(0)})$通过逐步左乘一系列的初等下三角矩阵$L_k(k = 1, \cdots, n-1)$，最终变为矩阵$(A^{(n-1)}, b^{(n-1)})$, 即将原方程组化为一个同解的上三角方程组$A^{(n-1)}x = b^{(n-1)}$.

事实上$A^{(0)} \triangleq A$,

$$A^{(1)} \triangleq \begin{pmatrix} a_{11}^{(0)} & a_{12}^{(0)} & \cdots & a_{1n}^{(0)} \\ 0 & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} \\ \vdots & \vdots & & \vdots \\ 0 & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} \end{pmatrix}, \quad L_1 \triangleq \begin{pmatrix} 1 & & & \\ -l_{21} & 1 & & \\ -l_{31} & & \ddots & \\ \vdots & & & \ddots & \\ -l_{n1} & & & & 1 \end{pmatrix},$$

则 $A^{(1)} = L_1 A^{(0)}$.

高斯消去法的矩阵形式

一般地,

$$A^{(k-1)} \triangleq \begin{pmatrix} a_{11}^{(0)} & a_{12}^{(0)} & \cdots & \cdots & a_{1n}^{(0)} \\ & a_{22}^{(1)} & \cdots & \cdots & a_{2n}^{(1)} \\ & & \ddots & & \vdots \\ & & a_{kk}^{(k-1)} & \cdots & a_{kn}^{(k-1)} \\ & & \vdots & & \vdots \\ & & a_{nk}^{(k-1)} & \cdots & a_{nn}^{(k-1)} \end{pmatrix} L_k \triangleq \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & -l_{k+1,k} & 1 & & \\ & & -l_{k+2,k} & & 1 & \\ & & \vdots & & & \ddots & \\ & & -l_{nk} & & & & 1 \end{pmatrix}$$

$$A^{(k)} \triangleq \begin{pmatrix} a_{11}^{(0)} & a_{12}^{(0)} & \cdots & a_{1,k+1} & \cdots & a_{1n}^{(0)} \\ & a_{22}^{(1)} & \cdots & a_{2,k+1} & \cdots & a_{2n}^{(1)} \\ & & \ddots & \vdots & & \vdots \\ & & a_{kk}^{(k-1)} & a_{k,k+1}^{(k-1)} & \cdots & a_{kn}^{(k-1)} \\ & & & a_{k+1,k+1}^{(k)} & \cdots & a_{k+1,n}^{(k)} \\ & & & \vdots & & \vdots \\ & & & a_{nk}^{(k)} & \cdots & a_{nn}^{(k)} \end{pmatrix}.$$

则有 $A^{(k)} = L_k A^{(k-1)} = L_k L_{k-1} \cdots L_2 L_1 A^{(0)}$.

高斯消去法的矩阵形式

一般地，

$$A^{(k-1)} \triangleq \begin{pmatrix} a_{11}^{(0)} & a_{12}^{(0)} & \cdots & \cdots & a_{1n}^{(0)} \\ & a_{22}^{(1)} & \cdots & \cdots & a_{2n}^{(1)} \\ & & \ddots & & \vdots \\ & & a_{kk}^{(k-1)} & \cdots & a_{kn}^{(k-1)} \\ & & \vdots & & \vdots \\ & & a_{nk}^{(k-1)} & \cdots & a_{nn}^{(k-1)} \end{pmatrix} L_k \triangleq \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & -l_{k+1,k} & 1 & & \\ & & -l_{k+2,k} & & 1 & \\ & & \vdots & & & \ddots & \\ & & -l_{nk} & & & & 1 \end{pmatrix}$$

$$A^{(k)} \triangleq \begin{pmatrix} a_{11}^{(0)} & a_{12}^{(0)} & \cdots & a_{1,k+1} & \cdots & a_{1n}^{(0)} \\ & a_{22}^{(1)} & \cdots & a_{2,k+1} & \cdots & a_{2n}^{(1)} \\ & & \ddots & \vdots & & \vdots \\ & & a_{kk}^{(k-1)} & a_{k,k+1}^{(k-1)} & \cdots & a_{kn}^{(k-1)} \\ & & & a_{k+1,k+1}^{(k)} & \cdots & a_{k+1,n}^{(k)} \\ & & & \vdots & & \vdots \\ & & & a_{nk}^{(k)} & \cdots & a_{nn}^{(k)} \end{pmatrix}.$$

则有 $A^{(k)} = L_k A^{(k-1)} = L_k L_{k-1} \cdots L_2 L_1 A^{(0)}$.

由于消元过程进行了$n-1$步，故

$$A^{(n-1)} = L_{n-1} A^{(n-2)} = \cdots = L_{n-1} L_{n-2} \cdots L_2 L_1 A^{(0)}$$

高斯消去法的矩阵形式

$$A^{(n-1)} \triangleq \begin{pmatrix} a_{11}^{(0)} & a_{12}^{(0)} & \cdots & a_{1n}^{(0)} \\ & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} \\ & & \ddots & \vdots \\ & & & a_{nn}^{(n-1)} \end{pmatrix}$$

同理 $b^{(n-1)} = L_{n-1}L_{n-2}\cdots L_2L_1b^{(0)}.$

注意到矩阵L_k的逆矩阵为

$$L_k^{-1} \triangleq \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & l_{k+1,k} & 1 & & \\ & & \vdots & & \ddots & \\ & & l_{nk} & & & 1 \end{pmatrix}$$

高斯消去法的矩阵形式

$$A^{(n-1)} \triangleq \begin{pmatrix} a_{11}^{(0)} & a_{12}^{(0)} & \cdots & a_{1n}^{(0)} \\ & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} \\ & & \ddots & \vdots \\ & & & a_{nn}^{(n-1)} \end{pmatrix}$$

同理　$b^{(n-1)} = L_{n-1}L_{n-2}\cdots L_2L_1b^{(0)}$.

注意到矩阵L_k的逆矩阵为

$$L_k^{-1} \triangleq \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & l_{k+1,k} & 1 & & \\ & & \vdots & & \ddots & \\ & & l_{nk} & & & 1 \end{pmatrix}$$

则

有$A = A^{(0)} = L_1^{-1}L_2^{-1}\cdots L_{n-1}^{-1}A^{(n-1)} = LU$,

$b = b^{(0)} = Lb^{(n-1)}$.

其中

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

$$U \triangleq A^{(n-1)} = \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ & u_{22} & \cdots & u_{2n} \\ & & \ddots & \\ & & & u_{nn} \end{pmatrix}.$$

Theorem

设A为n阶矩阵，如果A的顺序主子式$D_k \neq 0$ $(k = 1, 2, \cdots, n)$，则A可以唯一的分解成一个单位下三角矩阵L与一个上三角矩阵U的乘积，即$A = LU$.

Theorem

设A为n阶矩阵，如果A的顺序主子式$D_k \neq 0$ $(k = 1, 2, \cdots, n)$，则A可以唯一的分解成一个单位下三角矩阵L与一个上三角矩阵U的乘积，即$A = LU$.

证明 由高斯消去法的矩阵形式可知，$A = LU$分解的存在性. 下面证明分解的唯一性. 设矩阵A有两种分解

$$A = LU = L_1 U_1.$$

其中L, L_1是单位下三角阵，U, U_1是上三角阵. 由A可逆知L与U均可逆，故有 $L_1^{-1} L = U_1 U^{-1}$. 该式左端为单位下三角阵，右端为上三角阵. 因此该式左右两端必是单位阵. 于是$L_1 = L$, $U_1 = U$. $A = LU$称为**矩阵A的LU分解**.

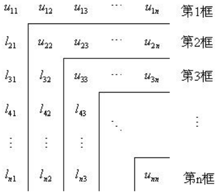
$A = LU$分解公式如下：

$$
\begin{cases}
u_{1j} = a_{1j}, \quad j = 1, 2, \cdots, n. \\
l_{i1} = a_{i1}/u_{11}, \quad i = 2, 3, \cdots, n. \\
u_{ij} = a_{ij} - \sum\limits_{k=1}^{i-1} l_{ik} u_{kj}, \quad i = 2, 3 \cdots, n, \quad j = i, i+1, \cdots, n. \\
l_{ij} = (a_{ij} - \sum\limits_{k=1}^{j-1} l_{ik} u_{kj})/u_{jj}, \; j = 2, 3, \cdots, n-1. \; i = j+1, j+2, \cdots, n.
\end{cases}
$$

$A=LU$分解公式如下：

$$
\begin{cases}
u_{1j}=a_{1j}, \quad j=1,2,\cdots,n. \\
l_{i1}=a_{i1}/u_{11}, \quad i=2,3,\cdots,n. \\
u_{ij}=a_{ij}-\sum\limits_{k=1}^{i-1} l_{ik}u_{kj}, \quad i=2,3\cdots,n, \quad j=i,i+1,\cdots,n. \\
l_{ij}=(a_{ij}-\sum\limits_{k=1}^{j-1} l_{ik}u_{kj})/u_{jj},\ j=2,3,\cdots,n-1.\ i=j+1,j+2,\cdots,n.
\end{cases}
$$

按先算矩阵U的第i行，再算矩阵L的第i列的顺序计算，计算流程按下图所示逐框进行.



矩阵$A=LU$分解的计算过程如下：

$$
\begin{aligned}
& u_{1j}=a_{1j}, \quad j=1,2,\cdots,n. \\
& l_{i1}=a_{i1}/a_{11}, \quad i=2,3,\cdots,n. \\
& for \quad i=2 \ to \ n-1 \ do \\
& u_{ii}:=a_{ii}-\sum_{k=1}^{i-1} l_{ik}u_{ki} \\
& \quad for \quad j=i+1 \ to \ n \ do \\
& \qquad u_{ij}:=a_{ij}-\sum_{k=1}^{i-1} l_{ik}u_{kj} \\
& \qquad l_{ji}:=(a_{ji}-\sum_{k=1}^{i-1} l_{jk}u_{ki})/u_{ii} \\
& \quad end \ do \\
& end \ do \\
& u_{nn}:=a_{nn}-\sum_{k=1}^{n-1} l_{nk}u_{kn}
\end{aligned}
$$

在计算机上实现矩阵$A=LU$分解时，为了减少存储量，将L的下三角部分的元素存放在A的下三角部分相应的位置上. （L 的对角元均等于1不用存储）. 将矩阵U的元存放于A的上三角部分位置，其编程时只需要将上述分解过程中l_{ij}, u_{ij} 改为相应的a_{ij} 即可.

矩阵的LU分解

在计算机上实现矩阵$A=LU$分解时，为了减少存储量，将L的下三角部分的元素存放在A的下三角部分相应的位置上.（L 的对角元均等于1不用存储）. 将矩阵U的元存放于A的上三角部分位置，其编程时只需要将上述分解过程中l_{ij}, u_{ij} 改为相应的a_{ij} 即可.

上面介绍的将矩阵分解为单位下三角阵L与上三角阵U的乘积$A=LU$，称为**杜里特尔(Doolittle)分解**. 同理矩阵A也可分解成下三角矩阵L与单位上三角矩阵U之积. 这种分解称为**克洛特(Crout)分解**.

例 分解矩阵$A = LU$, 其中

$$A = \begin{pmatrix} 4 & -2 & 0 & 4 \\ -2 & 2 & -3 & 1 \\ 0 & -3 & 13 & -7 \\ 4 & 1 & -7 & 23 \end{pmatrix}.$$

例 分解矩阵$A=LU$, 其中

$$A=\begin{pmatrix} 4 & -2 & 0 & 4 \\ -2 & 2 & -3 & 1 \\ 0 & -3 & 13 & -7 \\ 4 & 1 & -7 & 23 \end{pmatrix}.$$

解 按计算公式得

$$A=LU=\begin{pmatrix} 1 & 0 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 & 0 \\ 0 & -3 & 1 & 0 \\ 1 & 3 & \frac{1}{2} & 1 \end{pmatrix}\begin{pmatrix} 4 & -2 & 0 & 4 \\ 0 & 1 & -3 & 3 \\ 0 & 0 & 4 & 2 \\ 0 & 0 & 0 & 9 \end{pmatrix}.$$

下面介绍利用矩阵$A=LU$分解求解线性方程组$Ax=b$.

$Ax=LUx=b$等价于下面两个三角方程组

$$Ly=b, \qquad Ux=y.$$

先从$Ly=b$解出y，再从$Ux=y$ 解出x.

解方程组$Ly=b$. $Ly=b$的第一个方程为$y_1=b_1$，$Ly=b$的第i个方程为

$$\sum_{j=1}^{i-1} l_{ij}y_j + y_i = b_i\ , \quad i=2,3,\cdots,n.$$

所以,

$$\begin{cases} y_1 = b_1, \\ y_i = b_i - \sum\limits_{j=1}^{i-1} l_{ij}y_j, \quad i=2,3,\cdots,n. \end{cases}$$

矩阵的LU分解

从y的计算公式可以看到，可不必先将A分解后再求解$Ly=b$. 计算y与A的分解可以同时进行，**具体做法是对A的增广矩阵$\bar{A}=(A,b)$进行LU分解，分解结束后，$\bar{A}$的第$n+1$列位置上的元即是y的元.**

从y的计算公式可以看到，可不必先将A分解后再求解$Ly=b$. 计算y与A的分解可以同时进行，**具体做法是对A的增广矩阵$\bar{A}=(A,b)$进行LU分解，分解结束后，$\bar{A}$的第$n+1$列位置上的元即是y的元.**

求得y后，再求解上三角方程组$Ux=y$，同高斯消去法的回代过程，可得

$$\begin{cases} x_n = y_n/u_{nn} \ , \\ x_i = (y_i - \sum\limits_{j=i+1}^{n} u_{ij}x_j)/u_{ii}, \quad i=n-1,n-2,\cdots,2,1. \end{cases}$$

用$A=LU$ 分解解方程$Ax=b$ 乘除法的运算量与高斯消去法的乘除法运算次数相同.

例 2.4 用$A=LU$分解求解线性方程组$Ax=b$, 其中

$$A=\begin{pmatrix} 9 & 18 & 9 & -27 \\ 18 & 45 & 0 & -45 \\ 9 & 0 & 126 & 9 \\ -27 & -45 & 9 & 135 \end{pmatrix},\quad b=\begin{pmatrix} 1 \\ 2 \\ 16 \\ 8 \end{pmatrix}.$$

矩阵的LU分解

例 2.4 用$A=LU$分解求解线性方程组$Ax=b$, 其中

$$A=\begin{pmatrix} 9 & 18 & 9 & -27 \\ 18 & 45 & 0 & -45 \\ 9 & 0 & 126 & 9 \\ -27 & -45 & 9 & 135 \end{pmatrix},\quad b=\begin{pmatrix} 1 \\ 2 \\ 16 \\ 8 \end{pmatrix}.$$

解 将向量b作为$\bar{A}$的第5列，按紧凑格式计算，得

9	18	9	-27	1
2	9	-18	9	0
1	-2	81	54	15
-3	1	$\frac{2}{3}$	9	1

从$Ux=y$，解得$x=(\frac{1}{9},\frac{1}{9},\frac{1}{9},\frac{1}{9})^T$.

Theorem

(对称矩阵的三角分解) 设A是n阶对称矩阵，若A的各阶顺序主子式均不等于零，则A可以唯一的分解为

$$A = LDL^T.$$

Theorem

(对称矩阵的三角分解) 设A是n阶对称矩阵，若A的各阶顺序主子式均不等于零，则A可以唯一的分解为

$$A = LDL^T.$$

Theorem

*(对称正定矩阵的乔列斯基（**Cholesky**）分解)* 设A是n阶对称正定矩阵，则存在一个可逆的下三角阵G，使

$$A = GG^T.$$

当限定G的对角元为正时，这种分解是唯一的.

由G是下三角阵($g_{ij}=0, j>i$)及A的对称性，故只考虑A的下三角部分的元素，即$i \geq j$ 的情况. 据$A=GG^T$，由矩阵乘法得

$$a_{ij}=\sum_{k=1}^{n} g_{ik}g_{jk}=\sum_{k=1}^{\min(i,j)} g_{ik}g_{jk}=\sum_{k=1}^{j} g_{ik}g_{jk}=\sum_{k=1}^{j-1} g_{ik}g_{jk}+g_{ij}g_{jj}\ .$$

由G是下三角阵($g_{ij}=0, j>i$)及A的对称性，故只考虑A的下三角部分的元素，即$i \geq j$ 的情况. 据$A=GG^T$，由矩阵乘法得

$$a_{ij}=\sum_{k=1}^{n} g_{ik}g_{jk}=\sum_{k=1}^{\min(i,j)} g_{ik}g_{jk}=\sum_{k=1}^{j} g_{ik}g_{jk}=\sum_{k=1}^{j-1} g_{ik}g_{jk}+g_{ij}g_{jj}\ .$$

当$i=j$时，$a_{jj}=\sum_{k=1}^{j-1} g_{jk}^2+g_{jj}^2\ .$

所以，$g_{jj}=(a_{jj}-\sum_{k=1}^{j-1} g_{jk}^2)^{\frac{1}{2}}.\quad j=1,2,\cdots,n\ .$

由G是下三角阵($g_{ij}=0, j>i$)及A的对称性，故只考虑A的下三角部分的元素，即$i\geq j$ 的情况. 据$A=GG^T$，由矩阵乘法得

$$a_{ij}=\sum_{k=1}^{n} g_{ik}g_{jk}=\sum_{k=1}^{\min(i,j)} g_{ik}g_{jk}=\sum_{k=1}^{j} g_{ik}g_{jk}=\sum_{k=1}^{j-1} g_{ik}g_{jk}+g_{ij}g_{jj}\ .$$

当$i=j$时，$a_{jj}=\sum\limits_{k=1}^{j-1} g_{jk}^2+g_{jj}^2\ .$

所以，$g_{jj}=(a_{jj}-\sum\limits_{k=1}^{j-1} g_{jk}^2)^{\frac{1}{2}}.\quad j=1,2,\cdots,n\ .$

当$i>j$时，

$$g_{ij}=(a_{ij}-\sum_{k=1}^{j-1} g_{ik}g_{jk})/g_{jj}\ ,\quad j=1,2,\cdots,n-1.\quad i=j+1,j+2,\cdots,n\ .$$

当$j=1$时，$g_{i1}=a_{i1}/g_{11},\quad i=2,3,\cdots,n.$

$A = GG^T$分解的算法组织

$$
\begin{aligned}
& g_{11} := \sqrt{a_{11}} \\
& g_{i1} := a_{i1}/g_{11}\ , \quad i = 2, 3, \cdots, n. \\
& for \quad j = 2\ \ to\ \ n-1\ \ do \\
& \quad g_{jj} := (a_{jj} - \sum_{k=1}^{j-1} g_{jk}^2)^{\frac{1}{2}} \\
& \quad for \quad i = j+1\ \ to\ \ n\ \ do \\
& \quad\quad g_{ij} := (a_{ij} - \sum_{k=1}^{j-1} g_{ik} g_{jk})/g_{jj} \\
& \quad end\ \ do \\
& end\ \ do \\
& g_{nn} := (a_{nn} - \sum_{k=1}^{n-1} g_{nk}^2)^{\frac{1}{2}}
\end{aligned}
$$

平方根法

用$A = GG^T$分解求解方程租$Ax = b$的方法称为平方根法.

$$Ax = b \Longrightarrow GG^T x = b \Longrightarrow \begin{cases} Gy = b \Longrightarrow \text{解出}\, y, \\ G^T x = y \Longrightarrow \text{解出}\, x. \end{cases}$$

由$Gy = b$知

$$\begin{cases} g_{11} y_1 = b_1, \\ g_{i1} y_1 + g_{i2} y_2 + \cdots + g_{ii} y_i = b_i \ , \quad i = 2, 3, \cdots, n. \end{cases}$$

由此可得

$$\begin{cases} y_1 = b_1 / g_{11}, \\ y_i = \left(b_i - \sum\limits_{k=1}^{i-1} g_{ik} y_k \right) / g_{ii} \ , \quad i = 2, 3, \cdots, n. \end{cases}$$

然后从$G^T x = y$, 解出x.

例 用平方根法求解方程组

$$\bar{A} = \begin{pmatrix} 9 & 18 & 9 & -27 & 1 \\ 18 & 45 & 0 & -45 & 2 \\ 9 & 0 & 126 & 9 & 16 \\ -27 & -45 & 9 & 135 & 8 \end{pmatrix}$$

例 用平方根法求解方程组

$$\bar{A} = \begin{pmatrix} 9 & 18 & 9 & -27 & 1 \\ 18 & 45 & 0 & -45 & 2 \\ 9 & 0 & 126 & 9 & 16 \\ -27 & -45 & 9 & 135 & 8 \end{pmatrix}$$

解 $A = GG^T$, 其中

$$G = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 6 & 3 & 0 & 0 \\ 3 & -6 & 9 & 0 \\ -9 & 3 & 6 & 3 \end{pmatrix}, \ y = \begin{pmatrix} \frac{1}{3} \\ 0 \\ \frac{5}{3} \\ \frac{1}{3} \end{pmatrix}.$$

因此,

$$\begin{cases} Gy = b \Longrightarrow y = (\frac{1}{3}, 0, \frac{5}{3}, \frac{1}{3})^T. \\ G^T x = y \Longrightarrow x = (\frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9})^T. \end{cases}$$

矩阵的Cholesky分解需要作$\frac{1}{6}n^3-\frac{1}{6}n$个乘除法与$n$个开方运算，然后解两个三角方程组需要作$n^2$个乘除法. 故用平方根法求解$Ax=b$. 共需要作$\frac{1}{6}n^3+n^2-\frac{1}{6}n$个乘除法和$n$个开方运算，但$n$个开方运算需要耗费较多的机器时间.

矩阵的Cholesky分解需要作$\frac{1}{6}n^3-\frac{1}{6}n$个乘除法与$n$个开方运算，然后解两个三角方程组需要作$n^2$个乘除法. 故用平方根法求解$Ax=b$. 共需要作$\frac{1}{6}n^3+n^2-\frac{1}{6}n$个乘除法和$n$个开方运算，但$n$个开方运算需要耗费较多的机器时间.

为了避免平方根法的开方运算，对A作$A=LDL^T$分解.
由$A=LU=LDD^{-1}U=LDL^T$. 知$L^T=D^{-1}U$. 根据杜里特尔分解式有

$$\begin{cases} d_i = a_{ii} - \sum_{k=1}^{i-1} l_{ik}^2 d_k, \quad i=1,2,\cdots,n. \\ l_{ij} = (a_{ij} - \sum_{k=1}^{j-1} l_{ik} d_k l_{jk})/d_j, \quad j=1,2,\cdots,n-1. \quad i=j+1,j+2,\cdots \end{cases}$$

改进平方根法

用$A = LDL^T$分解求解方程组$Ax = b$，称为改进平方根法.

$$Ax = b \Longrightarrow LDL^T x = b \Longrightarrow \begin{cases} Ly = b \Longrightarrow \text{解出} y, \\ L^T x = D^{-1} y \Longrightarrow \text{解出} x. \end{cases}$$

用$A = LDL^T$分解求解方程组$Ax = b$，称为改进平方根法.

$$Ax = b \Longrightarrow LDL^T x = b \Longrightarrow \begin{cases} Ly = b \Longrightarrow \text{解出} y, \\ L^T x = D^{-1} y \Longrightarrow \text{解出} x. \end{cases}$$

由$Ly = b$ 得$y_i = b_i - \sum_{k=1}^{i-1} l_{ik} y_k, \quad i = 2, 3, \cdots, n.$

改进平方根法

用$A = LDL^T$分解求解方程组$Ax = b$，称为改进平方根法.

$$Ax = b \Longrightarrow LDL^T x = b \Longrightarrow \begin{cases} Ly = b \Longrightarrow \text{解出} y, \\ L^T x = D^{-1} y \Longrightarrow \text{解出} x. \end{cases}$$

由$Ly = b$ 得$y_i = b_i - \sum_{k=1}^{i-1} l_{ik} y_k, \quad i = 2, 3, \cdots, n.$

由$L^T x = D^{-1} z$ 得

$$\begin{cases} x_n = y_n / d_n \\ x_i = \frac{y_i}{d_i} - \sum\limits_{k=i+1}^{n} l_{ki} x_k, \quad i = n-1, n-2, \cdots, 2, 1. \end{cases}$$

用$A=LDL^T$分解求解方程组$Ax=b$，称为改进平方根法.

$$Ax=b \Longrightarrow LDL^Tx=b \Longrightarrow \begin{cases} Ly=b \Longrightarrow \text{解出}y, \\ L^Tx=D^{-1}y \Longrightarrow \text{解出}x. \end{cases}$$

由$Ly=b$ 得$y_i=b_i-\sum_{k=1}^{i-1} l_{ik}y_k, \quad i=2,3,\cdots,n.$

由$L^Tx=D^{-1}z$ 得

$$\begin{cases} x_n=y_n/d_n \\ x_i=\frac{y_i}{d_i}-\sum\limits_{k=i+1}^{n} l_{ki}x_k, \quad i=n-1,n-2,\cdots,2,1. \end{cases}$$

用改进平方根法解方程共用$\frac{1}{6}n^3+\frac{3}{2}n^2-\frac{2}{3}n$个乘除法，它比平方根法多了$\frac{1}{2}n^2-\frac{1}{2}n$个乘除法，少了$n$ 个开方运算.

求解三对角方程组$Ax = d$，即

$$\begin{pmatrix} b_1 & c_1 & & & \\ a_2 & b_2 & c_2 & & \\ & \ddots & \ddots & \ddots & \\ & & a_{n-1} & b_{n-1} & c_{n-1} \\ & & & a_n & b_n \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_{n-1} \\ d_n \end{pmatrix}.$$

求解三对角方程组的追赶法

求解三对角方程组$Ax = d$，即

$$\begin{pmatrix} b_1 & c_1 & & & \\ a_2 & b_2 & c_2 & & \\ & \ddots & \ddots & \ddots & \\ & & a_{n-1} & b_{n-1} & c_{n-1} \\ & & & a_n & b_n \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_{n-1} \\ d_n \end{pmatrix}.$$

对于三对角矩阵A，则有如下形式的杜里特尔分解.

$$A = LU = \begin{pmatrix} 1 & & & & \mathbf{0} \\ l_2 & 1 & & & \\ & \ddots & \ddots & & \\ \mathbf{0} & & l_{n-1} & 1 & \\ & & & l_n & 1 \end{pmatrix} \begin{pmatrix} u_1 & c_1 & & & \mathbf{0} \\ & u_2 & c_2 & & \\ & & \ddots & \ddots & \\ \mathbf{0} & & & u_{n-1} & c_{n-1} \\ & & & & u_n \end{pmatrix}$$

用矩阵$A = LU$分解求解三对角方程组称为**追赶法**.

$$\begin{cases} u_1 = b_1, \\ l_i = a_i / u_{i-1}, \quad i = 2, 3, \cdots, n. \\ u_i = b_i - l_i c_{i-1}, \quad i = 2, 3, \cdots, n. \end{cases}$$

即

$$\text{求解 } Ax = d \Longrightarrow LUx = d \Longrightarrow \begin{cases} Ly = d, \\ Ux = y. \end{cases}$$

用矩阵$A = LU$分解求解三对角方程组称为**追赶法**.

$$\begin{cases} u_1 = b_1, \\ l_i = a_i/u_{i-1}, \quad i = 2, 3, \cdots, n. \\ u_i = b_i - l_i c_{i-1}, \quad i = 2, 3, \cdots, n. \end{cases}$$

即

$$求解\ Ax = d \Longrightarrow LUx = d \Longrightarrow \begin{cases} Ly = d, \\ Ux = y. \end{cases}$$

由$Ly = d$，得

$$\begin{cases} y_1 = d_1, \\ y_i = d_i - l_i y_{i-1}, \quad i = 2, 3, \cdots, n. \end{cases}$$

用矩阵$A=LU$分解求解三对角方程组称为**追赶法**.

$$\begin{cases} u_1 = b_1, \\ l_i = a_i/u_{i-1}, \quad i=2,3,\cdots,n. \\ u_i = b_i - l_i c_{i-1}, \quad i=2,3,\cdots,n. \end{cases}$$

即

$$\text{求解 } Ax=d \Longrightarrow LUx=d \Longrightarrow \begin{cases} Ly=d, \\ Ux=y. \end{cases}$$

由$Ly=d$，得

$$\begin{cases} y_1 = d_1, \\ y_i = d_i - l_i y_{i-1}, \quad i=2,3,\cdots,n. \end{cases}$$

由$Ux=y$，得

$$\begin{cases} x_n = y_n/u_n, \\ x_i = (y_i - c_i x_{i+1})/u_i\ , \quad i=n-1,n-2,\cdots,1. \end{cases}$$

求解三对角方程组的追赶法

求解三对角方程$Ax = d$的追赶法的算法组织.

$$
\begin{aligned}
& u_1 := b_1, \quad y_1 := d_1 \\
& \mathit{for} \quad i = 2 \ \mathit{to} \ n \ \mathit{do} \\
& \quad l_i := a_i / u_{i-1} \\
& \quad u_i := b_i - l_i c_{i-1} \\
& \quad y_i := d_i - l_i y_{i-1} \\
& \mathit{end} \ \mathit{do} \\
& x_n := y_n / u_n \\
& \mathit{for} \quad i = n-1 \ \mathit{to} \ 1 \ \mathit{step} \ -1 \ \mathit{do} \\
& \quad x_i := (y_i - c_i x_{i+1}) / u_i \\
& \mathit{end} \ \mathit{do}
\end{aligned}
$$

求解三对角方程组的追赶法

求解三对角方程$Ax = d$的追赶法的算法组织.

$$
\begin{aligned}
& u_1 := b_1, \quad y_1 := d_1 \\
& for \quad i = 2\ to\ n\ do \\
& \quad l_i := a_i / u_{i-1} \\
& \quad u_i := b_i - l_i c_{i-1} \\
& \quad y_i := d_i - l_i y_{i-1} \\
& end\ do \\
& x_n := y_n / u_n \\
& for \quad i = n-1\ to\ 1\ step\ -1\ do \\
& \quad x_i := (y_i - c_i x_{i+1}) / u_i \\
& end\ do
\end{aligned}
$$

计算l_i, u_i, y_i的过程称为“**追**”的过程，计算x_i的过程称为“**赶**”的过程. 追赶法的乘除法的运算次数仅为$5n-4$. 由于方程组系数矩阵A是严格对角占优阵，保证了追赶法能顺利进行，且计算过程稳定.

例 2.6 用追赶法求解线性方程组$Ax = d$. 其中

$$A = \begin{pmatrix} 1 & 2 & 0 & 0 & 0 \\ 2 & 3 & 1 & 0 & 0 \\ 0 & -3 & 4 & 2 & 0 \\ 0 & 0 & 4 & 7 & 1 \\ 0 & 0 & 0 & -5 & 6 \end{pmatrix}, \quad d = \begin{pmatrix} 5 \\ 9 \\ 2 \\ 19 \\ -4 \end{pmatrix}.$$

例 2.6 用追赶法求解线性方程组$Ax=d$. 其中

$$A=\begin{pmatrix}1&2&0&0&0\\2&3&1&0&0\\0&-3&4&2&0\\0&0&4&7&1\\0&0&0&-5&6\end{pmatrix},\quad d=\begin{pmatrix}5\\9\\2\\19\\-4\end{pmatrix}.$$

解 追赶法求解方程组有

$$L=\begin{pmatrix}1&0&0&0&0\\2&1&0&0&0\\0&3&1&0&0\\0&0&4&1&0\\0&0&0&5&1\end{pmatrix},U=\begin{pmatrix}1&2&0&0&0\\0&-1&1&0&0\\0&0&1&2&0\\0&0&0&-1&1\\0&0&0&0&1\end{pmatrix},y=\begin{pmatrix}5\\-1\\5\\-1\\1\end{pmatrix}.$$

从 $Ly=d$ 解得 $y=(5,-1,5,-1,1)^T$, 从 $Ux=y$ 解得 $x=(1,2,1,2,1)^T$.

Definition

（向量范数） 设$f(x) = \|x\|$是定义在R^n上的非负实值函数，若$\|x\|$满足以下三条：

（1）$\|x\| \geq 0$, $\|x\| = 0 \Leftrightarrow x = 0$;

（2）$\forall \alpha \in R$, $\|\alpha x\| = |\alpha| \, \|x\|$;

（3）$\|x + y\| \leq \|x\| + \|y\|$;（三角不等式）

则称$\|x\|$是R^n上的向量范数.

Definition

（**向量范数**） 设$f(x)=\|x\|$是定义在R^n上的非负实值函数，若$\|x\|$满足以下三条：

（1）$\|x\|\geq 0,\ \|x\|=0\Leftrightarrow x=0;$

（2）$\forall\alpha\in R,\ \|\alpha x\|=|\alpha|\,\|x\|;$

（3）$\|x+y\|\leq\|x\|+\|y\|;$（三角不等式）

则称$\|x\|$是R^n上的向量范数.

三种常用的向量范数：

$$\|x\|_1=\sum_{i=1}^{n}|x_i|\text{（向量的1-范数）；}$$

$$\|x\|_2=\left(\sum_{i=1}^{n}x_i^2\right)^{\frac{1}{2}}\text{（向量的2-范数）；}$$

$$\|x\|_\infty=\max_{1\leq i\leq n}|x_i|\text{（向量的}\infty\text{-范数）．}$$

例如向量 $x=(2,-5,3)^T$ 的三种范数分别为：

$$\|x\|_1=10,\qquad \|x\|_2=\sqrt{38},\qquad \|x\|_\infty=5.$$

利用向量范数定义x的近似向量$\tilde{x}$的绝对误差与相对误差如下：
绝对误差：$\|\Delta x\|=\|x-\tilde{x}\|$，相对误差：$\frac{\|\Delta x\|}{\|x\|}=\frac{\|x-\tilde{x}\|}{\|x\|}$.

例如向量 $x=(2,-5,3)^T$ 的三种范数分别为：

$$\|x\|_1=10,\qquad \|x\|_2=\sqrt{38},\qquad \|x\|_\infty=5.$$

利用向量范数定义x的近似向量$\tilde{x}$的绝对误差与相对误差如下：
绝对误差：$\|\Delta x\|=\|x-\tilde{x}\|$，相对误差：$\frac{\|\Delta x\|}{\|x\|}=\frac{\|x-\tilde{x}\|}{\|x\|}$.

Theorem

（向量范数的连续性） 设$\|x\|$是R^n上任一种向量范数，则$\|x\|$是x的连续函数.

例如向量 $x=(2,-5,3)^T$ 的三种范数分别为：

$$\|x\|_1=10,\qquad \|x\|_2=\sqrt{38},\qquad \|x\|_\infty=5.$$

利用向量范数定义x的近似向量$\tilde{x}$的绝对误差与相对误差如下：
绝对误差：$\|\Delta x\|=\|x-\tilde{x}\|$, 相对误差：$\frac{\|\Delta x\|}{\|x\|}=\frac{\|x-\tilde{x}\|}{\|x\|}$.

Theorem

（向量范数的连续性） 设$\|x\|$是R^n上任一种向量范数，则$\|x\|$是x的连续函数.

Theorem

（向量范数的等价性） 设$\|x\|_p$和$\|x\|_q$是R^n上任意两种向量范数，则存在常数$c_1>0$，$c_2>0$，使得对任意的$x\in R^n$,有

$$c_1\|x\|_q\le\|x\|_p\le c_2\|x\|_q.$$

常用的几个等价关系如下：

$$\begin{cases} \|x\|_\infty \le \|x\|_1 \le n\|x\|_\infty; \\ \frac{1}{\sqrt{n}}\|x\|_1 \le \|x\|_2 \le \|x\|_1; \\ \frac{1}{\sqrt{n}}\|x\|_2 \le \|x\|_\infty \le \|x\|_2. \end{cases}$$

向量范数与矩阵范数

常用的几个等价关系如下：

$$\begin{cases} \|x\|_\infty \le \|x\|_1 \le n\|x\|_\infty; \\ \frac{1}{\sqrt{n}}\|x\|_1 \le \|x\|_2 \le \|x\|_1; \\ \frac{1}{\sqrt{n}}\|x\|_2 \le \|x\|_\infty \le \|x\|_2. \end{cases}$$

Definition

（矩阵范数） 设$f(A)=\|A\|$是定义在$R^{n\times n}$上的非负实值函数，若$\|A\|$满足以下四条：

（1）$\|A\| \ge 0$, $\|A\| = 0 \Leftrightarrow A = O$;

（2）$\forall \alpha \in R$, $\|\alpha A\| = |\alpha|\,\|A\|$;

（3）$\|A+B\| \le \|A\| + \|B\|$;（三角不等式）

（4）$\|AB\| \le \|A\|\|B\|$.

则称$\|A\|$是$R^{n\times n}$上的矩阵范数.

向量范数与矩阵范数

Definition

若矩阵范数$\|A\|$与向量范数$\|x\|$满足

$$\|Ax\| \le \|A\| \, \|x\|, \quad \forall x \in R, A \in R^{n\times n}$$

则称矩阵范数$\|A\|$与向量范数$\|x\|$是相容的或协调的.

Definition

若矩阵范数$\|A\|$与向量范数$\|x\|$满足

$$\|Ax\| \leq \|A\| \, \|x\|, \quad \forall x \in R, A \in R^{n\times n}$$

则称矩阵范数$\|A\|$与向量范数$\|x\|$是相容的或协调的.

Theorem

（矩阵的算子范数） 设$x \in R^n$, $A \in R^{n\times n}$, $\|x\|_p$ $(p = 1, 2, \infty)$ 是给定的一种向量范数，相应的定义一个矩阵的非负函数

$$\|A\|_p = \max_{x\neq 0} \frac{\|Ax\|_p}{\|x\|_p} = \max_{\|x\|_p=1} \|Ax\|_p \ .$$

则，$\|A\|_p$是一种矩阵范数，称为A的算子范数，且满足相容性条件$\|Ax\|_p \leq \|A\|_p\|x\|_p$.

证明 （1）显然，$\|A\|_p = \max\limits_{\|x\|_p=1} \|Ax\|_p \geq 0$,

$$\|A\|_p = \max_{x\neq 0} \frac{\|Ax\|_p}{\|x\|_p} = 0 \Leftrightarrow \|Ax\|_p = 0, \forall x \neq 0 \Leftrightarrow Ax = 0, \forall x \neq 0$$

所以，$\|A\|_p = 0 \Leftrightarrow A = O$.

证明 （1）显然，$\|A\|_p = \max\limits_{\|x\|_p=1} \|Ax\|_p \geq 0$,

$$\|A\|_p = \max_{x\neq 0} \frac{\|Ax\|_p}{\|x\|_p} = 0 \Leftrightarrow \|Ax\|_p = 0, \forall x \neq 0 \Leftrightarrow Ax = 0, \forall x \neq 0$$

所以，$\|A\|_p = 0 \Leftrightarrow A = O$.

（2） $\forall \alpha \in R$,

$$\|\alpha A\|_p = \max_{\|x\|_p=1} \|\alpha Ax\|_p = |\alpha| \max_{\|x\|_p=1} \|Ax\|_p = |\alpha| \, \|A\|_p \ .$$

证明　(1) 显然，$\|A\|_p = \max_{\|x\|_p=1} \|Ax\|_p \geq 0$,

$$\|A\|_p = \max_{x\neq 0} \frac{\|Ax\|_p}{\|x\|_p} = 0 \Leftrightarrow \|Ax\|_p = 0, \forall x \neq 0 \Leftrightarrow Ax = 0, \forall x \neq 0$$

所以，$\|A\|_p = 0 \Leftrightarrow A = O$.

(2)　$\forall \alpha \in R$,

$$\|\alpha A\|_p = \max_{\|x\|_p=1} \|\alpha Ax\|_p = |\alpha| \max_{\|x\|_p=1} \|Ax\|_p = |\alpha| \, \|A\|_p \ .$$

(3)　证三角不等式前先证相容性. 对于$\forall x \neq 0$,

$$\frac{\|Ax\|_p}{\|x\|_p} \leq \max_{x\neq 0} \frac{\|Ax\|_p}{\|x\|_p} = \|A\|_p \Rightarrow \|Ax\|_p \leq \|A\|_p \|x\|_p \ .$$

利用相容性证三角不等式，

$$
\begin{aligned}
\|A+B\|_p = \max_{\|x\|_p=1}\|(A+B)x\|_p = \max_{\|x\|_p=1}\|Ax+Bx\|_p \\
\le \max_{\|x\|_p=1}\|Ax\|_p + \max_{\|x\|_p=1}\|Bx\|_p \\
\le \max_{\|x\|_p=1}\|A\|_p\|x\|_p + \max_{\|x\|_p=1}\|B\|_p\,\|x\|_p = \|A\|_p + \|B\|_p\ .
\end{aligned}
$$

利用相容性证三角不等式，

$$\begin{aligned}\|A+B\|_p = \max_{\|x\|_p=1}\|(A+B)x\|_p &= \max_{\|x\|_p=1}\|Ax+Bx\|_p \\ &\le \max_{\|x\|_p=1}\|Ax\|_p + \max_{\|x\|_p=1}\|Bx\|_p \\ \le \max_{\|x\|_p=1}\|A\|_p\|x\|_p + \max_{\|x\|_p=1}\|B\|_p\,\|x\|_p &= \|A\|_p + \|B\|_p\ .\end{aligned}$$

（4）由相容性有

$$\|ABx\|_p \le \|A\|_p\|Bx\|_p \le \|A\|_p\|B\|_p\,\|x\|_p\ .$$

$\forall x\neq 0$，则

$$\frac{\|ABx\|_p}{\|x\|_p} \le \|A\|_p\|B\|_p\ .$$

故

$$\|AB\|_p = \max_{x\neq 0}\frac{\|ABx\|_p}{\|x\|_p} \le \|A\|_p\|B\|_p\ .$$

Theorem

设$x \in R^n, A \in R^{n\times n}$，按公式*(2.3.3)*确定的对应于向量的三种范数$\|x\|_p$ $(p=1,2,\infty)$的矩阵范数分别为

$$(1) \quad \|A\|_1 = \max_{1\le j\le n}\left\{\sum_{i=1}^{n}|a_{ij}|\right\} \quad \text{（1-范数或列范数）；}$$

$$(2) \quad \|A\|_2 = \sqrt{\lambda_{\max}(A^TA)} \quad \text{（2-范数或谱范数）；}$$

$$(3) \quad \|A\|_\infty = \max_{1\le i\le n}\left\{\sum_{j=1}^{n}|a_{ij}|\right\} \quad \text{（}\infty\text{ 范数或行范数）．}$$

其中$\lambda_{\max}(A^TA)$表示矩阵A^TA的最大特征值.

证明 （1）$A=(a_1,a_2,\cdots,a_n)$，则

$$\|Ax\|_1=\|x_1a_1+x_2a_2+\cdots+x_na_n\|_1\le|x_1|\,\|a_1\|_1+\cdots+|x_n|\,\|a_n\|_1$$
$$\le(|x_1|+|x_2|+\cdots+|x_n|)\max_{1\le j\le n}\{\|a_j\|_1\}$$
$$=\|x\|_1\max_{1\le j\le n}\{\|a_j\|_1\}.$$

而

$$\|A\|_1=\max_{\|x\|_1=1}\|Ax\|_1\le\max_{1\le j\le n}\{\|a_j\|_1\}=\max_{1\le j\le n}\left\{\sum_{i=1}^n|a_{ij}|\right\}.$$

证明　（1）$A=(a_1,a_2,\cdots,a_n)$，则

$$\begin{aligned}\|Ax\|_1 &= \|x_1a_1+x_2a_2+\cdots+x_na_n\|_1 \le |x_1|\,\|a_1\|_1+\cdots+|x_n|\,\|a_n\|_1\\ &\le (|x_1|+|x_2|+\cdots+|x_n|)\max_{1\le j\le n}\{\|a_j\|_1\}\\ &= \|x\|_1\max_{1\le j\le n}\{\|a_j\|_1\}.\end{aligned}$$

而

$$\|A\|_1=\max_{\|x\|_1=1}\|Ax\|_1\le\max_{1\le j\le n}\{\|a_j\|_1\}=\max_{1\le j\le n}\left\{\sum_{i=1}^n|a_{ij}|\right\}.$$

下面说明存在一个向量x，$\|x\|_1=1$，使 $\|Ax\|_1=\max_{1\le j\le n}\{\|a_j\|_1\}$.

不妨假定$\max_{1\le j\le n}\{\|a_j\|_1\}=\|a_k\|_1$.取$x=e_k$，显然$\|e_k\|=1$，且

$$\|Ae_k\|_1=\|a_k\|_1=\max_{1\le j\le n}\{\|a_j\|_1\}=\max_{1\le j\le n}\left\{\sum_{i=1}^n|a_{ij}|\right\}.$$

于是

$$\|A\|_1 = \max_{\|x\|_1=1} \|Ax\|_1 = \max_{1\le j\le n} \left\{ \sum_{i=1}^{n} |a_{ij}| \right\} .$$

于是

$$\|A\|_1 = \max_{\|x\|_1=1} \|Ax\|_1 = \max_{1\le j\le n} \left\{ \sum_{i=1}^{n} |a_{ij}| \right\} .$$

（2）注意到A^TA是实对称且至少是半正定矩阵，由线性代数理论知A^TA有完全的标准正交特征向量系$u_i(i=1,2,\cdots,n)$.
取$u_i(i=1,2,\cdots,n)$作为R^n的一组基底. $\forall x\in R, x\neq 0$, 则

$$x = \sum_{i=1}^{n} c_i u_i \ .$$

于是

$$\frac{\|Ax\|_2^2}{\|x\|_2^2} = \frac{(Ax,Ax)}{(x,x)} = \frac{(A^TAx,x)}{(x,x)} = \frac{\sum\limits_{i=1}^{n} c_i^2\lambda_i}{\sum\limits_{i=1}^{n} c_i^2} \le \lambda_1 \ .$$

另一方面取$x=u_1$, 则上式等号成立.

例 2.7 计算矩阵$A=\begin{pmatrix} 1 & -2 \\ 3 & 4 \end{pmatrix}$的三种范数.

例 2.7 计算矩阵$A=\begin{pmatrix} 1 & -2 \\ 3 & 4 \end{pmatrix}$的三种范数.

解

$$\|A\|_1=\max\{4,6\}=6,\quad \|A\|_\infty=\max\{3,7\}=7.$$

例 2.7 计算矩阵$A=\begin{pmatrix}1 & -2\\ 3 & 4\end{pmatrix}$的三种范数.

解

$$\|A\|_1=\max\{4,6\}=6,\quad \|A\|_\infty=\max\{3,7\}=7.$$

以下求$\|A\|_2$.

$$A^TA=\begin{pmatrix}1 & 3\\ -2 & 4\end{pmatrix}\begin{pmatrix}1 & -2\\ 3 & 4\end{pmatrix}=\begin{pmatrix}10 & 10\\ 10 & 20\end{pmatrix},$$

$$|\lambda I-A^TA|=\begin{vmatrix}\lambda-10 & -10\\ -10 & \lambda-20\end{vmatrix}=\lambda^2-30\lambda+100=0.$$

解特征方程, 得$\lambda=15\pm5\sqrt{5}$. 所以, $\lambda_{\max}(A^TA)=15+5\sqrt{5}$.
故$\|A\|_2=\sqrt{15+5\sqrt{5}}$.

还有一种常用的矩阵范数是F($Frobenius$)范数:

$$\|A\|_F = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^2}\ .$$

向量范数与矩阵范数

还有一种常用的矩阵范数是$F(Frobenius)$范数:

$$\|A\|_F=\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}^2}\ .$$

矩阵的F范数与向量的2-范数$\|x\|_2$是相容的，即

$$\|Ax\|_2\le\|A\|_F\|x\|_2\ .$$

事实上

$$\|Ax\|_2^2=\sum_{i=1}^{n}\left(\sum_{j=1}^{n}a_{ij}x_j\right)^2\le\sum_{i=1}^{n}\left(\sum_{j=1}^{n}a_{ij}^2\sum_{k=1}^{n}x_k^2\right)$$

$$=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}^2\bullet\sum_{k=1}^{n}x_k^2=\|A\|_F^2\ \|x\|_2^2\ .$$

上式两边开方即得结论.

矩阵范数也具有其等价性. 即$\|A\|_p$, $\|A\|_q$是$R^{n\times n}$ 中任意两种矩阵范数，则存在常数$c_1 > 0$, $c_2 > 0$, 使

$$c_1\|A\|_q \le \|A\|_p \le c_2\|A\|_q \ , \ \forall A \in R^{n\times n}$$

成立.

矩阵范数也具有其等价性. 即$\|A\|_p$, $\|A\|_q$是$R^{n\times n}$ 中任意两种矩阵范数，则存在常数$c_1>0$, $c_2>0$, 使

$$c_1\|A\|_q \le \|A\|_p \le c_2\|A\|_q \ , \ \forall A \in R^{n\times n}$$

成立.

Definition

设$A\in R^{n\times n}$, $\lambda_i\,(i=1,2,\cdots,n)$ 是A的特征值，则

$$\rho(A)=\max_{1\le i\le n}\{|\lambda_i|\}$$

称为矩阵A的谱半径.

Theorem

设$A \in R^{n\times n}$，则$\rho(A) \le \|A\|$. 即谱半径不超过A的任一种范数.

Theorem

设$A \in R^{n\times n}$，则$\rho(A) \le \|A\|$. 即谱半径不超过A的任一种范数.

证明 设λ是A的任一特征值，x是对应的特征向量，则

$$Ax = \lambda x.$$

于是， $|\lambda|\,\|x\| = \|\lambda x\| = \|Ax\| \le \|A\|\,\|x\|$. 注意到$\|x\| \ne 0$, 则得

$$|\lambda| \le \|A\| .$$

因而，$\rho(A) \le \|A\|$.

Theorem

设$\|B\| < 1$, 则$I - B$ 是可逆矩阵，且有

$$\|(I-B)^{-1}\| \leq \frac{1}{1-\|B\|}\ .$$

Theorem

设$\|B\|<1$, 则$I-B$ 是可逆矩阵，且有

$$\|(I-B)^{-1}\| \le \frac{1}{1-\|B\|} \ .$$

证明　用反证法证$I-B$ 可逆. 假定$I-B$ 不可逆，则方程组$(I-B)x=0$ 有非零解. 于是

$$\begin{aligned} 0 &= \|0\| = \|(I-B)x\| = \|x-Bx\| \ge \|x\|-\|Bx\| \ge \|x\|-\|B\|\,\|x\| \\ &= (1-\|B\|)\|x\|. \end{aligned}$$

得到矛盾.
由$I-B$ 可逆，则有

$$\begin{aligned} 1 &= \|I\| = \|(I-B)(I-B)^{-1}\| = \|(I-B)^{-1}-B(I-B)^{-1}\| \\ &\ge \|(I-B)^{-1}\|-\|B(I-B)^{-1}\| \ge \|(I-B)^{-1}\|-\|B\|\,\|(I-B)^{-1}\| \\ &= (1-\|B\|)\,\|(I-B)^{-1}\| \ . \end{aligned}$$

从而有　$\|(I-B)^{-1}\| \le \frac{1}{1-\|B\|}$.

Theorem

设线性方程组$Ax=b$ $(|A|\neq 0,\ b\neq 0)$ 的系数矩阵A和右端项b有微小的扰动ΔA、Δb，扰动后的方程组为

$$(A-\Delta A)(x-\Delta x)=b-\Delta b\ .$$

当$\|A^{-1}\|\,\|\Delta A\|<1$时，则方程组$Ax=b$ 近似解$\tilde{x}$ 的相对误差估计式为

$$\frac{\|\Delta x\|}{\|x\|}\leq\frac{\|A\|\,\|A^{-1}\|}{1-\|A\|\,\|A^{-1}\|\,\frac{\|\Delta A\|}{\|A\|}}\left(\frac{\|\Delta b\|}{\|b\|}+\frac{\|\Delta A\|}{\|A\|}\right)\ .$$

Theorem

设线性方程组$Ax=b$ ($|A|\neq 0$, $b\neq 0$) 的系数矩阵A和右端项b有微小的扰动ΔA、Δb，扰动后的方程组为

$$(A-\Delta A)(x-\Delta x)=b-\Delta b\ .$$

当$\|A^{-1}\|\,\|\Delta A\|<1$时，则方程组$Ax=b$ 近似解$\tilde{x}$ 的相对误差估计式为

$$\frac{\|\Delta x\|}{\|x\|}\leq\frac{\|A\|\,\|A^{-1}\|}{1-\|A\|\,\|A^{-1}\|\,\frac{\|\Delta A\|}{\|A\|}}\left(\frac{\|\Delta b\|}{\|b\|}+\frac{\|\Delta A\|}{\|A\|}\right)\ .$$

当$\frac{\|\Delta A\|}{\|A\|}$ 较小时，近似解的相对误差$\frac{\|\Delta x\|}{\|x\|}$ 约为$\frac{\|\Delta A\|}{\|A\|}$ 与$\frac{\|\Delta b\|}{\|b\|}$ 之和的$\|A\|\,\|A^{-1}\|$倍. $\|A\|\,\|A^{-1}\|$越小，近似解的相对误差就越小. $\|A\|\,\|A^{-1}\|$ 反映了原始数据对解的影响.

证明 由$(A-\Delta A)(x-\Delta x)=b-\Delta b$和$Ax=b$得

$$(A-\Delta A)\Delta x=\Delta b-\Delta A\,x\ ,$$

即

$$(I-A^{-1}\Delta A)\Delta x=A^{-1}(\Delta b-\Delta A\,x)\ .$$

由定理(2.3.6) 知$I-A^{-1}\Delta A$ 可逆，且

$$\|(I-A^{-1}\Delta A)^{-1}\|\leq\frac{1}{1-\|A^{-1}\Delta A\|}\leq\frac{1}{1-\|A^{-1}\|\,\|\Delta A\|}\ .$$

知

$$\Delta x=(I-A^{-1}\Delta A)^{-1}A^{-1}(\Delta b-\Delta Ax)\ .$$

两边取范数，有

$$
\begin{aligned}
\|\Delta x\| &\le \|(I - A^{-1}\Delta A)^{-1}\| \, \|A^{-1}\| \, (\|\Delta b\| + \|\Delta A\| \, \|x\|) \\
&\le \frac{\|A^{-1}\|}{1 - \|A^{-1}\| \, \|\Delta A\|} \left(\|\Delta b\| + \|\Delta A\| \, \|x\|\right) .
\end{aligned}
$$

上式两端除以$\|x\|$, 得

$$
\frac{\|\Delta x\|}{\|x\|} \le \frac{\|A^{-1}\|}{1 - \|A^{-1}\| \, \|\Delta A\|} \left(\frac{\|\Delta b\|}{\|x\|} + \|\Delta A\| \right)
$$

两边取范数，有

$$\begin{aligned}\|\Delta x\| &\le \|(I - A^{-1}\Delta A)^{-1}\|\,\|A^{-1}\|\,(\|\Delta b\| + \|\Delta A\|\,\|x\|)\\ &\le \frac{\|A^{-1}\|}{1-\|A^{-1}\|\,\|\Delta A\|}(\|\Delta b\| + \|\Delta A\|\,\|x\|)\,.\end{aligned}$$

上式两端除以$\|x\|$, 得

$$\begin{aligned}\frac{\|\Delta x\|}{\|x\|} &\le \frac{\|A^{-1}\|}{1-\|A^{-1}\|\,\|\Delta A\|}\left(\frac{\|\Delta b\|}{\|x\|} + \|\Delta A\|\right)\\ &= \frac{\|A\|\,\|A^{-1}\|}{1-\|A\|\,\|A^{-1}\|\,\frac{\|\Delta A\|}{\|A\|}}\left(\frac{\|\Delta b\|}{\|A\|\,\|x\|} + \frac{\|\Delta A\|}{\|A\|}\right)\,.\end{aligned}$$

注意到$\|b\| = \|Ax\| \le \|A\|\|x\|$. 则得

$$\frac{\|\Delta x\|}{\|x\|} \le \frac{\|A\|\,\|A^{-1}\|}{1-\|A\|\,\|A^{-1}\|\,\frac{\|\Delta A\|}{\|A\|}}\left(\frac{\|\Delta b\|}{\|b\|} + \frac{\|\Delta A\|}{\|A\|}\right)\,.$$

Definition

设A 是非奇异矩阵，数

$$\mathrm{Cond}(A) \triangleq \|A\| \, \|A^{-1}\|$$

称为矩阵A 的条件数. 条件数与矩阵的范数有关，也将条件数记为

$$\mathrm{Cond}_p(A) \triangleq \|A\|_p \, \|A^{-1}\|_p \ (p = 1, 2, \infty).$$

常用的矩阵条件数为

（1） $\mathrm{Cond}_\infty(A) = \|A\|_\infty \, \|A^{-}\|_\infty$;

（2） $\mathrm{Cond}_2(A) = \|A\|_2 \, \|A^{-1}\|_2 = \sqrt{\dfrac{\lambda_{\max}(A^TA)}{\lambda_{\min}(AA^T)}}$.

当A是对称矩阵时$\mathrm{Cond}_2(A) = \frac{|\lambda_1|}{|\lambda_n|}$，其中$\lambda_1$，$\lambda_n$ 分别是矩阵A的绝对值最大和绝对值最小的特征值.

矩阵的条件数刻画了方程组的性态，**条件数大的矩阵称为“病态”矩阵，相应的方程组称为“病态”方程组；条件数小的矩阵称为“良态”矩阵，相应的方程组称为“良态”方程组.**

矩阵的条件数刻画了方程组的性态，**条件数大的矩阵称为“病态”矩阵，相应的方程组称为“病态”方程组；条件数小的矩阵称为“良态”矩阵，相应的方程组称为“良态”方程组.**

例2.9 设有线性方程组

$$\begin{pmatrix} 1 & 10^5 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 10^5 \\ 2 \end{pmatrix}.$$

矩阵的条件数刻画了方程组的性态，**条件数大的矩阵称为“病态”矩阵，相应的方程组称为“病态”方程组；条件数小的矩阵称为“良态”矩阵，相应的方程组称为“良态”方程组.**

例2.9 设有线性方程组

$$\begin{pmatrix} 1 & 10^5 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 10^5 \\ 2 \end{pmatrix}.$$

则 $$A^{-1} = \frac{1}{10^5 - 1} \begin{pmatrix} -1 & 10^5 \\ 1 & -1 \end{pmatrix},$$

$$\|A\|_\infty = 10^5 + 1, \quad \|A^{-1}\|_\infty = \frac{10^5+1}{10^5-1}, \quad \mathrm{Cond}_\infty(A) > 10^5 \gg 1.$$

例2.10 n 阶Hilbert阵, 随着矩阵H_n阶数n的增高，H_n的条件数急剧增大，此时方程组$H_n x = b$是严重的病态方程组.

例2.10 n 阶Hilbert阵, 随着矩阵H_n阶数n的增高，H_n的条件数急剧增大，此时方程组$H_nx=b$是严重的病态方程组.
三阶Hilbert方程组为例说明其病态性.

$$\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \frac{11}{6} \\ \frac{13}{12} \\ \frac{47}{60} \end{pmatrix},$$

$$H_3^{-1} = \begin{pmatrix} 9 & -36 & 30 \\ -36 & 192 & -180 \\ 30 & -180 & 180 \end{pmatrix}.$$

$\|H_3\|_\infty = \dfrac{11}{6}$, $\|H_3^{-1}\|_\infty = 408$, $\mathrm{Cond}_\infty(H_3) = \|H_3\|_\infty \|H_3^{-1}\|_\infty = 748$

若将方程组中的分数舍入到两位小数后求解，则求得的近似解为
$\tilde{x} = (-6.22\,, 38.25\,, -33.65)^T$, 与真解$x=(1,1,1)^T$ 相差甚远.

舍入误差对解的影响

例2.10 n 阶Hilbert阵, 随着矩阵H_n阶数n的增高，H_n的条件数急剧增大，此时方程组$H_nx=b$是严重的病态方程组.
三阶Hilbert方程组为例说明其病态性.

$$\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \frac{11}{6} \\ \frac{13}{12} \\ \frac{47}{60} \end{pmatrix},$$

$$H_3^{-1} = \begin{pmatrix} 9 & -36 & 30 \\ -36 & 192 & -180 \\ 30 & -180 & 180 \end{pmatrix}.$$

$\|H_3\|_\infty = \dfrac{11}{6}$, $\quad \|H_3^{-1}\|_\infty = 408$, $\quad \mathrm{Cond}_\infty(H_3) = \|H_3\|_\infty \|H_3^{-1}\|_\infty = 748$

若将方程组中的分数舍入到两位小数后求解，则求得的近似解为$\tilde{x} = (-6.22\,, 38.25\,, -33.65)^T$，与真解$x=(1,1,1)^T$ 相差甚远.
$\mathrm{Cond}_\infty(H_6) = 2.9\times 10^7$, $\mathrm{Cond}_\infty(H_7) = 9.85\times 10^8$

病态方程组的特征/判别

条件数计算很复杂，在实际中难以应用，往往具有下列现象之一的方程组可能是病态方程组.

（1）系数矩阵的行或列近似线性相关（如例2.8）；

病态方程组的特征/判别

条件数计算很复杂，在实际中难以应用，往往具有下列现象之一的方程组可能是病态方程组.

（1）系数矩阵的行或列近似线性相关（如例2.8）；

（2）系数矩阵的元素数量级相差悬殊（如例2.9，例2.10）；

病态方程组的特征/判别

条件数计算很复杂，在实际中难以应用，往往具有下列现象之一的方程组可能是病态方程组.

（1）系数矩阵的行或列近似线性相关（如例2.8）；

（2）系数矩阵的元素数量级相差悬殊（如例2.9，例2.10）；

（3）解对系数矩阵元素的变化比较敏感；

病态方程组的特征/判别

条件数计算很复杂，在实际中难以应用，往往具有下列现象之一的方程组可能是病态方程组.

（1）系数矩阵的行或列近似线性相关（如例2.8）；

（2）系数矩阵的元素数量级相差悬殊（如例2.9，例2.10）；

（3）解对系数矩阵元素的变化比较敏感；

（4）在主元素消去法求解过程中，出现量级很小的主元素；

病态方程组的特征/判别

条件数计算很复杂，在实际中难以应用，往往具有下列现象之一的方程组可能是病态方程组.

（1）系数矩阵的行或列近似线性相关（如例2.8）；

（2）系数矩阵的元素数量级相差悬殊（如例2.9，例2.10）；

（3）解对系数矩阵元素的变化比较敏感；

（4）在主元素消去法求解过程中，出现量级很小的主元素；

（5）求出的解与预期的解相差较大；

病态方程组的特征/判别

条件数计算很复杂，在实际中难以应用，往往具有下列现象之一的方程组可能是病态方程组.

（1）系数矩阵的行或列近似线性相关（如例2.8）；

（2）系数矩阵的元素数量级相差悬殊（如例2.9，例2.10）；

（3）解对系数矩阵元素的变化比较敏感；

（4）在主元素消去法求解过程中，出现量级很小的主元素；

（5）求出的解与预期的解相差较大；

求解病态方程组应采取的措施:

（1）采用高精度（如双精度）计算，以减少舍入误差的影响；

病态方程组的特征/判别

条件数计算很复杂，在实际中难以应用，往往具有下列现象之一的方程组可能是病态方程组.

（1）系数矩阵的行或列近似线性相关（如例2.8）；

（2）系数矩阵的元素数量级相差悬殊（如例2.9，例2.10）；

（3）解对系数矩阵元素的变化比较敏感；

（4）在主元素消去法求解过程中，出现量级很小的主元素；

（5）求出的解与预期的解相差较大；

求解病态方程组应采取的措施:

（1）采用高精度（如双精度）计算，以减少舍入误差的影响；

（2）采用稳定性好的算法（如全主元高斯消去法）；

病态方程组的特征/判别

条件数计算很复杂，在实际中难以应用，往往具有下列现象之一的方程组可能是病态方程组.

（1）系数矩阵的行或列近似线性相关（如例2.8）；

（2）系数矩阵的元素数量级相差悬殊（如例2.9，例2.10）；

（3）解对系数矩阵元素的变化比较敏感；

（4）在主元素消去法求解过程中，出现量级很小的主元素；

（5）求出的解与预期的解相差较大；

求解病态方程组应采取的措施:

（1）采用高精度（如双精度）计算，以减少舍入误差的影响；

（2）采用稳定性好的算法（如全主元高斯消去法）；

（3）采取平衡措施；所谓平衡措施是将系数矩阵的各行除以该行绝对值最大的元素.

例如对例2.9 的系数矩阵平衡后变为$A=\begin{pmatrix} 10^{-5} & 1 \\ 1 & 1 \end{pmatrix}$, $A^{-1}=\frac{1}{1-10^{-5}}\begin{pmatrix} -1 & 1 \\ 1 & -10^{-5} \end{pmatrix}$,
$\|A\|\infty=2$, $\|A^{-1}\|\infty=\frac{2}{1-10^{-5}}$, $\mathrm{Cond}_\infty(A)\approx 4$.

例如对例2.9 的系数矩阵平衡后变为$A=\begin{pmatrix} 10^{-5} & 1 \\ 1 & 1 \end{pmatrix}$, $A^{-1}=\frac{1}{1-10^{-5}}\begin{pmatrix} -1 & 1 \\ 1 & -10^{-5} \end{pmatrix}$,
$\|A\|\infty=2$, $\|A^{-1}\|\infty=\frac{2}{1-10^{-5}}$, $\mathrm{Cond}_{\infty}(A)\approx 4$.

（4）采用迭代改善技术；迭代改善技术适用于不是十分病态的线性方程组. 其实质是采用迭代的思想，对近似解不断修正，使得近似解逐渐逼近方程组的准确解.

求修正量Δx，使得$\tilde{x}+\Delta x$ 满足方程组$Ax=b$,
即求解方程组$A(\tilde{x}+\Delta\tilde{x})=b$. $A\Delta x=r=b-A\tilde{x}$.
r 称为**残向量**,求得Δx后，重复以上过程，直到求得满足精度要求的近似解.

迭代改善技术

迭代改善技术计算过程如下：

(1) 给定误差限$\epsilon > 0$，将A进行三角分解$A = LU$，求得方程组$Ax = b$的一个近似解$x^{(0)}$，$0 \Rightarrow k$；

(2) 用双精度计算$r^{(k)} = b - Ax^{(k)}$；

(3) 解方程组$Ly = r^{(k)}$， $U\Delta x = y$，设其解为$\Delta x^{(k)}$；

(4) 令$x^{(k+1)} = x^{(k)} + \Delta x^{(k)}$；

(5) 如果$\|\Delta x^{(k)}\| \leq \epsilon$，或者$\frac{\|\Delta x^{(k)}\|}{\|x^{(k+1)}\|} \leq \epsilon$，则取$x^{(k+1)}$作为方程组$Ax = b$ 的近似解，停；否则，$k + 1 \Rightarrow k$，转步（2）.