矩阵方程
约束解的迭代算法

张凯院 编著

Juzhen Fangcheng
Yueshujie de Diedai Suanfa

国防工業出版社
National Defense Industry Press

矩阵方程约束解的迭代算法

张凯院　编著

国防工业出版社

·北京·

内容简介

本书共分为7章,内容包括:预备知识,分块带状线性代数方程组的PE解法,线性矩阵方程的分组迭代解法和参数迭代解法,线性矩阵方程约束解的MCG算法,非线性矩阵方程约束解的双迭代算法,以及MCG算法的应用等.

本书内容新颖,反映了线性矩阵方程唯一解的某些迭代算法、线性矩阵方程约束解的MCG算法和非线性矩阵方程约束解的双迭代算法研究的最新进展,可作为高等院校理工科研究生和数学专业高年级本科生的教学用书,也可作为相关专业科研和技术人员的参考书.

图书在版编目(CIP)数据

矩阵方程约束解的迭代算法/张凯院编著.—北京:国防工业出版社,2015.5

ISBN 978-7-118-10055-6

Ⅰ.①矩… Ⅱ.①张… Ⅲ.①矩阵-线性方程-迭代计算 Ⅳ.①O241.6

中国版本图书馆CIP数据核字(2015)第081194号

※

国防工業出版社 出版发行

(北京市海淀区紫竹院南路23号 邮政编码100048)

国防工业出版社印刷厂印刷

新华书店经售

*

开本 787×1092 1/16 **印张** 15¾ **字数** 356千字

2015年5月第1版第1次印刷 **印数** 1—1500册 **定价** 36.00元

(本书如有印装错误,我社负责调换)

国防书店:(010)88540777 发行邮购:(010)88540776

发行传真:(010)88540755 发行业务:(010)88540717

前　言

理论上,可将线性矩阵方程(LME)转化为线性代数方程组进行求解计算.但是,当未知矩阵的阶数较高时,这种转化方式产生的线性代数方程组的阶数按照平方量级增长(如果未知矩阵是方阵的话),求解的计算量也就加大.因此,直接建立LME的求解算法是有意义的.

对于线性矩阵方程的唯一解问题,可以采用按行拉直的方法转化为线性代数方程组,建立古典迭代算法,并利用这种转化得到的线性代数方程组的系数矩阵具有特殊的分块结构,将古典迭代算法转化为分组迭代算法.这样,每一组迭代计算涉及的矩阵阶数与原线性矩阵方程中对应的矩阵阶数相同.当然,也可以针对某些特殊的线性矩阵方程,直接建立简单而有效的迭代算法.

约束矩阵方程问题就是在满足一定条件的矩阵集合中求矩阵方程的解.不同的矩阵方程或不同的约束条件都将导致不同的约束矩阵方程问题.对于多矩阵变量的矩阵方程,当解矩阵属于同一约束矩阵集合时,称为同类约束解;当解矩阵属于不同的约束矩阵集合时,称为异类约束解.约束解的类型有一般解、对称解、反对称解、中心对称解、中心反对称解、自反解、反自反解、双对称解、对称次反对称解、对称自反解、对称反自反解、同类约束解和异类约束解以及其他的子空间约束解.

基于求解线性代数方程组的共轭梯度法的思想方法,通过对下降方向和极小点的修正处理,可以建立求LME的某类约束解的新型迭代算法,称为修正共轭梯度法(MCG算法).MCG算法能够自动判断LME是否有某类约束解,不要求涉及的线性代数方程组(指LME的向量形式)的系数矩阵对称正定、可逆或者列满秩,但仍然具备有限步收敛性,理论上总是可行的.当LME没有某类约束解时,可以将求它的约束最小二乘解问题转化为求它的约束正规矩阵方程的约束解问题,针对后者建立求其约束解的MCG算法,就是求LME的约束最小二乘解的MCG算法.

采用MCG算法求由牛顿算法每一步迭代计算导出的关于校正矩阵的LME的约束解或者约束最小二乘解,可以建立求非线性矩阵方程的约束解的双迭代算法.通常的牛顿算法要求这些LME有唯一的某类约束解,这就需要对非线性矩阵方程的系数矩阵做一些附加限定.双迭代算法不要求这些LME有某类约束解,也就不需要对非线性矩阵方程的系数矩阵做任何附加限定,原因是MCG算法总是可行的.对于牛顿算法中的某一步,当导出的关于校正矩阵的LME没有某类约束解时,可用它的这类约束最小二乘解来代替.当然,求LME的约束最小二乘解比求LME的约束解的计算量要大,因为约束正规矩阵方程比原来的LME复杂得多.

本书分为7章,第1章简要介绍求解线性代数方程组的古典迭代方法、变分原理迭代方法和整体校正加速方法,矩阵的广义逆及其等价定义和应用,矩阵的直积运算与行拉直

向量，以及后续章节讨论涉及的特殊矩阵．第 2 章介绍求块三对角、块五对角和周期块三对角线性代数方程组唯一解的 PE 方法，并讨论 PE 方法的收敛性问题．第 3 章介绍求一类 Lyapunov 矩阵方程和一般线性矩阵方程唯一解的分组迭代方法，并讨论分组迭代方法的收敛性问题．第 4 章介绍求几类特殊 Lyapunov 矩阵方程唯一解的参数迭代方法，并讨论参数迭代方法的收敛性和最优参数选取问题．第 5 章介绍当 LME 有约束解，但约束解不一定唯一时，求其一组约束解的 MCG 算法；当 LME 没有约束解时，求其一组约束最小二乘解的 MCG 算法．第 6 章介绍 MCG 算法在求非线性矩阵方程约束解方面的应用，也就是求非线性矩阵方程约束解的双迭代算法. 第 7 章介绍 MCG 算法在计算广义逆矩阵和求解子空间约束矩阵方程(包括子空间约束线性代数方程组)方面的应用．

本书的部分内容取自作者编著的《数值代数》修订本(科学出版社 2010 年出版)，大部分资料来源于作者与指导的研究生近期发表的研究论文. 作者无意探究某些算法的历史根源，因为这是一项极其复杂的工作，只是通过一些个例来描述算法的构造原理或者应用场合．求解矩阵方程的迭代算法多种多样，书中只是介绍作者比较熟悉的部分迭代算法．如果将线性代数方程组看作线性矩阵方程的特例，再将一般解问题看作约束解问题的特例，本书的内容当属矩阵方程约束解的迭代算法，牵强地以此定为书名．

本书内容新颖，反映了线性矩阵方程唯一解某些迭代算法、线性矩阵方程约束解修正共轭梯度法和非线性矩阵方程约束解双迭代算法的最新进展；个例代表性强，覆盖面宽，表明了迭代算法的构造原理；结构紧凑，叙述简明，突出了矩阵分析方法. 阅读本书只需具备线性代数和计算方法的基本知识．

本书在撰写过程中，得到了西北工业大学理学院、教务处和研究生院等部门的大力支持. 多届研究生，如李书连、刘晓敏、朱寿升、牛婷婷、宋卫红、宁倩芝等，在数值计算与资料整理方面做了大量工作．特别是西北工业大学邓子辰教授和长安大学封建湖教授审阅了书稿，并提出了有益的修改建议，作者在此表示衷心的感谢．

作者由衷地感谢西北工业大学专著出版基金资助本书的出版．

由于作者水平有限，书中疏漏及不妥之处在所难免，敬请同行和读者批评指正．

作 者

2014 年 10 月于西北工业大学

符 号 说 明

符号	含义
$\mathbf{R}(\mathbf{C})$	实(复)数集合
$\mathbf{R}^n(\mathbf{C}^n)$	实(复)n 维向量集合
$\mathbf{R}^{m\times n}(\mathbf{C}^{m\times n})$	实(复)$m\times n$ 矩阵集合
$\mathbf{R}_r^{m\times n}(\mathbf{C}_r^{m\times n})$	秩为 r 的实(复)$m\times n$ 矩阵集合
$\mathrm{span}\{\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_m\}$	由元素 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_m$ 生成的子空间
$\mathbf{0}$	零向量或线性空间的零元素
$\boldsymbol{e}_i$	第 i 个分量为 1,其余分量为 0 的向量
$\boldsymbol{O}$	零矩阵
$\boldsymbol{I}$	单位矩阵
$\boldsymbol{E}_{ij}$	第 i 行 j 列元素为 1,其余元素为 0 的矩阵
$\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$	以 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 为对角元素的对角矩阵
$\det\boldsymbol{A}$	方阵 $\boldsymbol{A}$ 的行列式
$\mathrm{tr}\boldsymbol{A}$	方阵 $\boldsymbol{A}$ 的迹
$\rho(\boldsymbol{A})$	方阵 $\boldsymbol{A}$ 的谱半径
$\mathrm{adj}\boldsymbol{A}$	方阵 $\boldsymbol{A}$ 的伴随矩阵
$\mathrm{rank}\boldsymbol{A}$	矩阵 $\boldsymbol{A}$ 的秩
$R(\boldsymbol{A})$	矩阵 $\boldsymbol{A}$ 的值域
$N(\boldsymbol{A})$	矩阵 $\boldsymbol{A}$ 的零空间
$\mathrm{vec}(\boldsymbol{A})$	矩阵 $\boldsymbol{A}$ 按列拉直的列向量
$\overline{\mathrm{vec}}(\boldsymbol{A})$	矩阵 $\boldsymbol{A}$ 按行拉直的列向量
$\boldsymbol{A}^{\mathrm{T}}$	矩阵 $\boldsymbol{A}$ 的转置
$\boldsymbol{A}^{\mathrm{H}}$	矩阵 $\boldsymbol{A}$ 的共轭转置
$\boldsymbol{A}^{+}$	矩阵 $\boldsymbol{A}$ 的 Moore - Penrose 逆
$\boldsymbol{A}^{(1)}$	矩阵 $\boldsymbol{A}$ 的$\{1\}$ - 逆
$\boldsymbol{A}^{(1,3)}$	矩阵 $\boldsymbol{A}$ 的$\{1,3\}$ - 逆
$\boldsymbol{A}^{(1,4)}$	矩阵 $\boldsymbol{A}$ 的$\{1,4\}$ - 逆
$\boldsymbol{A}^{(d)}$	矩阵 $\boldsymbol{A}$ 的 Drazin 逆

$A\{1\}$	矩阵 $\boldsymbol{A}$ 的 $\{1\}$ - 逆的集合
$\boldsymbol{A}\otimes\boldsymbol{B}$	矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的直积
$[\boldsymbol{x},\boldsymbol{y}]$	元素 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 的内积
$\boldsymbol{x}\perp\boldsymbol{y}$	元素 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 正交
$\Vert\boldsymbol{x}\Vert_p$	向量 $\boldsymbol{x}$ 的 p - 范数
$\Vert\boldsymbol{x}\Vert_A$	向量 $\boldsymbol{x}$ 的椭圆范数
$\Vert\boldsymbol{A}\Vert_F$	矩阵 $\boldsymbol{A}$ 的 Frobenius 范数
$a\in S$	元素 a 属于集合 S
$a\notin S$	元素 a 不属于集合 S
$\boldsymbol{W}^{\perp}$	子空间 $\boldsymbol{W}$ 的正交补
$\dim V$	线性空间 V 的维数
$V_1\cap V_2$	集合 V_1 与 V_2 的交
$V_1\cup V_2$	集合 V_1 与 V_2 的并
V_1+V_2	子空间 V_1 与 V_2 的和
$V_1\oplus V_2$	子空间 V_1 与 V_2 的直和
$\partial f(\lambda)$	多项式 $f(\lambda)$ 的次数
$f(\lambda)\mid g(\lambda)$	多项式 $f(\lambda)$ 整除多项式 $g(\lambda)$
Ω_0	一般实矩阵集合
$\Omega_1,\mathbf{SR}^{n\times n}$	n 阶实对称矩阵集合
$\Omega_2,\mathbf{ASR}^{n\times n}$	n 阶实反对称矩阵集合
$\Omega_3,\mathbf{CSR}^{n\times n}$	n 阶中心对称矩阵集合
$\Omega_4,\mathbf{CASR}^{n\times n}$	n 阶中心反对称矩阵集合
$\Omega_5,\mathbf{R}_r^{n\times n}(\boldsymbol{P})$	关于对称正交矩阵 $\boldsymbol{P}$ 的 n 阶自反矩阵集合
$\Omega_6,\mathbf{R}_a^{n\times n}(\boldsymbol{P})$	关于对称正交矩阵 $\boldsymbol{P}$ 的 n 阶反自反矩阵集合
$\Omega_7,\mathbf{BSR}^{n\times n}$	n 阶双对称矩阵集合
$\Omega_8,\mathbf{SASR}^{n\times n}$	n 阶对称次反对称矩阵集合
$\Omega_9,\mathbf{SR}_r^{n\times n}(\boldsymbol{P})$	关于对称正交矩阵 $\boldsymbol{P}$ 的 n 阶对称自反矩阵集合
$\Omega_{10},\mathbf{SR}_a^{n\times n}(\boldsymbol{P})$	关于对称正交矩阵 $\boldsymbol{P}$ 的 n 阶对称反自反矩阵集合
$\mathbf{R}_r^{n\times n}(\boldsymbol{P}_1,\boldsymbol{P}_2)$	关于对称正交矩阵 $\boldsymbol{P}_1$、$\boldsymbol{P}_2$ 的 n 阶广义自反矩阵集合
$\mathbf{R}_a^{n\times n}(\boldsymbol{P}_1,\boldsymbol{P}_2)$	关于对称正交矩阵 $\boldsymbol{P}_1$、$\boldsymbol{P}_2$ 的 n 阶广义反自反矩阵集合
$(\boldsymbol{X}_1,\boldsymbol{X}_2)\in\Omega_{1-3}$	$\boldsymbol{X}_1\in\Omega_1$ 且 $\boldsymbol{X}_2\in\Omega_3$

目　录

第 1 章　预备知识 ……………………………………………… 1

1.1　古典迭代方法 ……………………………………………… 1

1.2　变分原理迭代方法 ……………………………………………… 4

1.2.1　最速下降法 ……………………………………………… 5

1.2.2　共轭梯度法 ……………………………………………… 5

1.3　整体校正加速方法 ……………………………………………… 6

1.3.1　线性代数方程组问题 ……………………………………………… 7

1.3.2　线性矩阵方程问题 ……………………………………………… 10

1.4　矩阵的广义逆 ……………………………………………… 14

1.4.1　广义逆矩阵的概念 ……………………………………………… 14

1.4.2　广义逆矩阵的等价定义 ……………………………………………… 15

1.4.3　广义逆矩阵的应用 ……………………………………………… 17

1.5　矩阵的直积运算与行拉直向量 ……………………………………………… 19

1.5.1　直积的概念 ……………………………………………… 19

1.5.2　线性矩阵方程的可解性 ……………………………………………… 20

1.6　特殊矩阵及其集合记号 ……………………………………………… 22

参考文献 ……………………………………………… 23

第 2 章　分块带状线性代数方程组的 PE 解法 ……………………………………………… 24

2.1　块三对角方程组的 PE 解法 ……………………………………………… 24

2.1.1　PE 方法 ……………………………………………… 24

2.1.2　PE 方法和 PE_k 方法的收敛性 ……………………………………………… 26

2.1.3　二次 PE 方法和二次 PE_k 方法的收敛性 ……………………………………………… 28

2.1.4　数值算例 ……………………………………………… 29

2.2　块五对角方程组的 PE 解法 ……………………………………………… 30

2.2.1　PE 方法 ……………………………………………… 30

2.2.2　PE 方法的收敛性 ……………………………………………… 33

2.2.3　PE_k 方法的收敛性 ……………………………………………… 36

2.2.4　二次 PE 方法和二次 PE_k 方法的收敛性 ……………………………………………… 37

2.2.5　数值算例 ……………………………………………… 38

2.3　周期块三对角方程组的 PE 解法 ……………………………………………… 40

2.3.1 PE 方法 …… 40
2.3.2 PE 方法和 PE_k 方法的收敛性 …… 41
2.3.3 二次 PE 方法的收敛性 …… 42
2.3.4 数值算例 …… 44
参考文献 …… 46

第 3 章 线性矩阵方程的分组迭代解法 …… 48

3.1 Lyapunov 矩阵方程的分组迭代解法 …… 48
3.1.1 Jacobi 和 JGS 迭代格式 …… 49
3.1.2 拟 JGS 迭代格式 …… 50
3.1.3 块 Jacobi 和块 JGS 迭代格式 …… 53
3.1.4 SOR、拟 SOR 和块 SOR 迭代格式 …… 57
3.2 一般线性矩阵方程的分组迭代解法 …… 62
3.2.1 Jacobi 和 JGS 迭代格式 …… 63
3.2.2 拟 JGS 迭代格式 …… 65
3.2.3 块 Jacobi 和块 JGS 迭代格式 …… 66
3.2.4 SOR、拟 SOR 和块 SOR 迭代格式 …… 68
参考文献 …… 73

第 4 章 线性矩阵方程的参数迭代解法 …… 74

4.1 矩阵方程 $\boldsymbol{AX}+\boldsymbol{XA}^{\mathrm{H}}=\boldsymbol{F}$ 的参数迭代解法 …… 74
4.2 矩阵方程 $\boldsymbol{AX}+\boldsymbol{XB}=\boldsymbol{F}$ 的参数迭代解法 …… 77
4.3 矩阵方程 $\boldsymbol{AXB}^{\mathrm{T}}+\boldsymbol{BXA}^{\mathrm{T}}=\boldsymbol{F}$ 的参数迭代解法 …… 82
4.4 矩阵方程 $\boldsymbol{AXB}+\boldsymbol{CXD}=\boldsymbol{F}$ 的参数迭代解法 …… 86
4.5 矩阵方程 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{X}+\boldsymbol{XA}+\boldsymbol{B}^{\mathrm{T}}\boldsymbol{XB}=\boldsymbol{C}$ 的参数迭代解法 …… 91
参考文献 …… 95

第 5 章 线性矩阵方程约束解的 MCG 算法 …… 96

5.1 求解线性代数方程组的 MCG 算法 …… 96
5.2 简单线性矩阵方程约束解的 MCG 算法 …… 98
5.2.1 迭代方法 …… 98
5.2.2 求一般解的算法收敛性分析 …… 101
5.3 一般线性矩阵方程约束解的 MCG 算法 …… 105
5.3.1 迭代方法 …… 105
5.3.2 求对称解的算法收敛性分析 …… 106
5.4 线性矩阵方程约束 Ls 解的 MCG 算法 …… 111
5.4.1 约束正规矩阵方程与迭代方法 …… 112
5.4.2 求中心对称 Ls 解的算法收敛性分析 …… 117
5.5 线性矩阵方程组约束解的 MCG 算法 …… 122

5.5.1 迭代方法 …… 122
5.5.2 求自反解的算法收敛性分析 …… 125
5.6 多变量线性矩阵方程组约束解的 MCG 算法 …… 131
5.6.1 迭代方法 …… 131
5.6.2 求一般解的算法收敛性分析 …… 135
5.7 线性矩阵方程异类约束解的 MCG 算法 …… 140
5.7.1 求一般异类约束解的迭代方法与收敛性结论 …… 141
5.7.2 求分组异类约束解的迭代方法与收敛性结论 …… 144
5.8 线性矩阵方程异类约束 Ls 解的 MCG 算法 …… 147
5.8.1 求约束 1-3 Ls 解的迭代方法与收敛性结论 …… 147
5.8.2 求约束 1-3-7 Ls 解的迭代方法与收敛性结论 …… 150
参考文献 …… 153

第 6 章 非线性矩阵方程约束解的双迭代算法 …… 156

6.1 高次多项式矩阵方程约束解的双迭代算法 …… 156
6.1.1 双迭代算法介绍 …… 156
6.1.2 求双变量矩阵方程的对称解 …… 159
6.1.3 求双变量矩阵方程的约束 1-9 解 …… 163
6.1.4 求双变量矩阵方程组的对称解 …… 169
6.2 含逆幂的矩阵方程约束解的双迭代算法 …… 173
6.2.1 求单变量矩阵方程的对称解 …… 173
6.2.2 求双变量矩阵方程组的约束 1-9 解 …… 177
6.2.3 求特殊结构的单变量矩阵方程的对称解 …… 184
6.3 含特殊逆幂的矩阵方程约束解的双迭代算法 …… 190
6.3.1 求含高次逆幂的单变量矩阵方程的对称解 …… 190
6.3.2 求含分数逆幂的单变量矩阵方程的约束解 …… 194
6.3.3 求含高次逆幂的双变量矩阵方程组的对称自反解 …… 203
参考文献 …… 209

第 7 章 MCG 算法的应用 …… 211

7.1 逆矩阵的迭代算法 …… 211
7.1.1 古典迭代方法 …… 211
7.1.2 Newton 迭代方法 …… 212
7.1.3 MCG 算法 …… 213
7.2 Moore-Penrose 逆的直接迭代算法 …… 214
7.3 Moore-Penrose 逆的 MCG 算法 …… 218
7.3.1 转化为求单变量线性矩阵方程的一般解 …… 218
7.3.2 转化为求单变量线性矩阵方程组的一般解 …… 219
7.3.3 转化为求双变量线性矩阵方程组的一般解 …… 221

7.3.4 数值算例 …… 222
7.4 Drazin 逆的 MCG 算法 …… 225
7.4.1 转化为求单变量线性矩阵方程的一般解 …… 225
7.4.2 转化为求单变量线性矩阵方程组的一般解 …… 226
7.4.3 数值算例 …… 226
7.5 矩阵方程子空间约束解的 MCG 算法 …… 228
7.5.1 求方程 $\boldsymbol{AXB}=\boldsymbol{D}$ 的子空间约束解 …… 229
7.5.2 求方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的子空间约束解 …… 232
7.5.3 求方程组 $\boldsymbol{Ax}+\boldsymbol{y}=\boldsymbol{b}$ 的子空间约束解 …… 233
7.5.4 求方程 $\boldsymbol{A}_1\boldsymbol{X}_1\boldsymbol{B}_1+\boldsymbol{A}_2\boldsymbol{X}_2\boldsymbol{B}_2=\boldsymbol{D}$ 的子空间约束解 …… 236
参考文献 …… 240

第 1 章　预 备 知 识

本章简要介绍求解线性代数方程组的古典迭代方法、变分原理迭代方法和整体校正加速方法，矩阵的广义逆及其等价定义和应用，矩阵的直积运算与行拉直向量，以及后续章节讨论的特殊矩阵.

1.1　古典迭代方法

本节介绍求解线性代数方程组

$$Ax = b \tag{1.1}$$

的古典迭代方法，其中 A 为 n 阶可逆矩阵，b 为 n 维列向量. 分裂 $A = M - N$，其中 M 为可逆矩阵，则方程组(1.1) 等价于

$$x = M^{-1}Nx + M^{-1}b.$$

令 $B = M^{-1}N, f = M^{-1}b$，上式可写为

$$x = Bx + f. \tag{1.2}$$

建立迭代格式

$$x^{(k+1)} = Bx^{(k)} + f \quad (k = 0,1,2,\cdots), \tag{1.3}$$

式中：B 为 n 阶矩阵，称为**迭代矩阵**；f 为与 A 和 b 有关的 n 维列向量；$x^{(0)}$ 为任意给定的 n 维列向量，称为**初始向量**.

如果向量序列 $\{x^{(k)}\}$ 收敛，即 $\lim\limits_{k\to\infty} x^{(k)} = x^*$，则 x^* 就是方程组(1.2) 的唯一解，从而也是方程组(1.1) 的唯一解.

就其分量而言，由于迭代格式(1.3) 中的 $x^{(k+1)}$ 线性依赖于 $x^{(k)}$，而不明显依赖于 $x^{(k-1)}, x^{(k-2)}, \cdots, x^{(0)}$，且迭代矩阵 B 保持不变，所以称格式(1.3) 为**一阶线性定常迭代格式**. 对方程组(1.1) 做不同的等价变形，将会得到不同的迭代格式(1.3)，其区别仅在于迭代矩阵 B 和向量 f. 下面介绍几种常用的迭代格式.

简单迭代格式：分裂 $A = I - (I - A)$，方程组(1.1) 等价于

$$x = (I - A)x + b. \tag{1.4}$$

令 $B = I - A, f = b$，即得式(1.2)，相应的格式(1.3) 称为**简单迭代格式**.

Jacobi 迭代格式：设 $A = (a_{ij})_{n\times n}$ 的主对角线元素 $a_{ii} \neq 0 (i = 1,2,\cdots,n)$，则对角矩阵 $D = \mathrm{diag}(a_{11}, a_{22}, \cdots, a_{nn})$ 可逆. 分裂 $A = D - (D - A)$，方程组(1.1) 等价于

$$x = D^{-1}(D - A)x + D^{-1}b. \tag{1.5}$$

令 $B_J = D^{-1}(D - A), f_J = D^{-1}b$，即得式(1.2)，相应的格式(1.3) 称为 **Jacobi 迭代格式.**

GS(Gauss - Seidel) 迭代格式：在方程组(1.1) 的等价形式(1.2) 中，分裂 $B = L + U$，其中 L 为严格下三角矩阵，U 为上三角矩阵，于是式(1.2) 等价于

$$x = (I - L)^{-1}Ux + (I - L)^{-1}f. \tag{1.6}$$

令 $B_{GS} = (I - L)^{-1}U, f_{GS} = (I - L)^{-1}f$,即得另一等价形式(1.2),相应的格式(1.3)称为**GS 迭代格式**,并可改写为

$$x^{(k+1)} = Lx^{(k+1)} + Ux^{(k)} + f \quad (k = 0,1,2,\cdots). \tag{1.7}$$

相对原来的格式(1.3)来说,格式(1.7)充分利用了向量 $x^{(k+1)}$ 的最新信息,即每求得 $x^{(k+1)}$ 的一个分量 $\xi_i^{(k+1)}$,就用它来代替 $\xi_i^{(k)}$ 进行后面的计算.

JGS 迭代格式:在 GS 迭代格式中,取原来的等价形式(1.2)为 $x = B_J x + f_J$ 时,相应的等价形式(1.6)为

$$x = B_{JGS}x + f_{JGS},$$

此时,相应的格式(1.3)称为 **JGS 迭代格式**. 由于矩阵 B_J 的主对角线元素全为0,所以相应的格式(1.7)的计算量会有所减少.

SOR 迭代格式(Jacobi 型):在方程组(1.1)的等价形式(1.2)中,取 B 为 Jacobi 格式的迭代矩阵 B_J,取 f 为 f_J,并分裂 $B_J = L + R$,其中 L 为严格下三角矩阵,R 为严格上三角矩阵. 那么,相应的 JGS 迭代格式为

$$x^{(k+1)} = Lx^{(k+1)} + Rx^{(k)} + f_J \quad (k = 0,1,2,\cdots).$$

对某个参数 ω,上式右端与 $x^{(k)}$ 作加权平均,可得新的迭代格式,即

$$x^{(k+1)} = (1 - \omega)x^{(k)} + \omega(Lx^{(k+1)} + Rx^{(k)} + f_J) \quad (k = 0,1,2,\cdots), \tag{1.8}$$

称之为 **SOR 迭代格式**,并称 ω 为**松弛因子**. 格式(1.8)可改写为格式(1.3)的形式,即

$$x^{(k+1)} = B_\omega x^{(k)} + f_\omega \quad (k = 0,1,2,\cdots),$$

式中

$$B_\omega = (I - \omega L)^{-1}[(1 - \omega)I + \omega R], \quad f_\omega = \omega(I - \omega L)^{-1}f_J.$$

当 $\omega = 1$ 时,SOR 格式就是 JGS 格式. 适当选取 ω,可使 SOR 格式的收敛速度优于 JGS 格式.

常用的向量范数与矩阵范数如下:设 $x = (\xi_1, \xi_2, \cdots, \xi_n)^T, A = (a_{ij})_{m\times n}$,那么向量 x 的 1 - 范数、2 - 范数和 ∞ - 范数依次为

$$\|x\|_1 = \sum_{i=1}^{n}|\xi_i|, \quad \|x\|_2 = \sqrt{|\xi_1|^2 + |\xi_2|^2 + \cdots + |\xi_n|^2}, \quad \|x\|_\infty = \max_i|\xi_i|.$$

设 B 是 Hermite 正定矩阵,那么向量 x 的椭圆范数为

$$\|x\|_B = \sqrt{x^H Bx}.$$

矩阵 A 的 1 - 范数、2 - 范数(谱范数)、∞ - 范数和 F - 范数依次为

$$\|A\|_1 = \max_{1\le j\le n}\sum_{i=1}^{m}|a_{ij}|, \quad \|A\|_2 = \sqrt{\lambda_1}\ (\lambda_1 \text{ 表示 } A^H A \text{ 的最大特征值}),$$

$$\|A\|_\infty = \max_{1\le i\le m}\sum_{j=1}^{n}|a_{ij}|, \quad \|A\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2}.$$

引理 1.1 对 $\mathbf{C}^{n\times n}$ 上的矩阵范数 $\|A\|_M$,存在 $\mathbf{C}^n$ 上的向量范数 $\|x\|_V$,使得

$$\|Ax\|_V \le \|A\|_M \cdot \|x\|_V \quad (\text{任意 } A \in \mathbf{C}^{n\times n}, \text{任意 } x \in \mathbf{C}^n),$$

称为矩阵范数 $\|A\|_M$ 与向量范数 $\|x\|_V$ 相容.

引理 1.2 设 $\|A\|_M$ 是 $\mathbf{C}^{n\times n}$ 上的矩阵范数,任取 $\mathbf{C}^n$ 中的非零列向量 y,则函数

$$\| x \|_V = \| x y^H \|_M \quad (\text{任意 } x \in \mathbf{C}^n)$$

是 $\mathbf{C}^n$ 上的向量范数,且矩阵范数 $\| A \|_M$ 与向量范数 $\| x \|_V$ 相容[7].

对于由格式(1.3)产生的向量序列 $\{x^{(k)}\}$,需要研究何时 $\lim\limits_{k\to\infty} x^{(k)} = x^*$;$k$ 固定时怎样估计误差 $\| x^{(k)} - x^* \|$ 的上界;k 多大时才能用 $x^{(k)}$ 作为 x^* 的近似值.

定理 1.1　对于任意的 $x^{(0)}$ 及 f,由格式(1.3)产生的向量序列 $\{x^{(k)}\}$ 收敛于 x^* 的充要条件是 $\rho(B) < 1$.

应用定理 1.1 判别向量序列 $\{x^{(k)}\}$ 的收敛性时,需要估计迭代矩阵 B 的按模最大特征值. 下面介绍可以避免这一计算的判断方法.

定理 1.2　对于格式(1.3),若有矩阵范数 $\| \cdot \|$ 使得 $\| B \| < 1$,则向量序列 $\{x^{(k)}\}$ 收敛于 x^*,且有

$$\| x^{(k)} - x^* \| \leqslant \frac{\| B \|^k}{1 - \| B \|}\left(\| x^{(0)} \| + \| f \| \right), \tag{1.9}$$

$$\| x^{(k)} - x^* \| \leqslant \frac{\| B \|}{1 - \| B \|} \| x^{(k)} - x^{(k-1)} \|, \tag{1.10}$$

式中的矩阵范数与向量范数相容.

式(1.9)给出了误差 $\| x^{(k)} - x^* \|$ 的上界估计方法. 当 k 充分大时,由 $\| B \| < 1$ 知 $\| B \|^k$ 充分小,只要 $\| B \|$ 不是很接近于1,式(1.9)的右端将充分小,从而可用 $x^{(k)}$ 代替 x^*. 式(1.10)给出了迭代次数 k 的估计方法. 设容许误差为 ε,即 $\| x^{(k)} - x^* \| \leqslant \varepsilon$ 时,可用 $x^{(k)}$ 代替 x^*,那么可由

$$\| x^{(k)} - x^{(k-1)} \| \leqslant \frac{1 - \| B \|}{\| B \|} \varepsilon$$

确定所需要的 k. 实际计算中,只要 $\| B \|$ 不是很接近于1,就可用 $\| x^{(k)} - x^{(k-1)} \| \leqslant \varepsilon$ 作为迭代过程的终止准则.

上面的定理未涉及迭代矩阵 B 的具体形式,其结论适用于一般的一阶线性定常迭代格式(1.3). 下面针对5种基本迭代格式的特殊情形,分别介绍它们的收敛性定理.

定理 1.3　设 $A = (a_{ij})_{n\times n}$ 实对称,且 $a_{ii} > 0 (i = 1,2,\cdots,n)$,则 Jacobi 格式收敛的充要条件是 A 和 $2D - A$ 都是正定矩阵.

定理 1.4　设方程组(1.2)中的矩阵 B 满足 $\| B \|_\infty < 1$,或者 $\| B \|_1 < 1$,则 GS 格式收敛.

定理 1.5　设 A 满足下列条件之一:

(1) A 按行(列)严格对角占优.

(2) A 按行(列)弱对角占优且不可约.

则 Jacobi 格式和 JGS 格式都收敛.

需要指出,定理 1.5 给出了 Jacobi 格式和 JGS 格式同时收敛的充分条件,但它们的收敛性态并非完全一致. 在方程组(1.1)中,对于一些矩阵 A,可能 Jacobi 格式和 JGS 格式同时收敛或同时发散(不收敛);但对于另一些矩阵 A,也可能其中一个格式收敛,而另外一个格式发散. 考虑以下的几个矩阵,即

$$A_1 = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 3 & -10 \\ 9 & -4 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 1 & 2 & -2 \\ 1 & 1 & 1 \\ 2 & 2 & 1 \end{bmatrix}, \quad A_4 = \begin{bmatrix} 2 & -1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & -2 \end{bmatrix}.$$

当 $\boldsymbol{A} = \boldsymbol{A}_1$ 时,因为 $\boldsymbol{A}$ 按行严格对角占优,根据定理 1.5,Jacobi 格式和 JGS 格式同时收敛. 当 $\boldsymbol{A} = \boldsymbol{A}_2$ 时,由于

$$\boldsymbol{B}_{\mathrm{J}} = \boldsymbol{D}^{-1}(\boldsymbol{D} - \boldsymbol{A}) = \begin{bmatrix} 0 & 10/3 \\ 9/4 & 0 \end{bmatrix}, \quad \rho(\boldsymbol{B}_{\mathrm{J}}) = \sqrt{7.5} > 1,$$

所以 Jacobi 格式发散;又

$$\boldsymbol{B}_{\mathrm{JGS}} = (\boldsymbol{I} - \boldsymbol{L})^{-1}\boldsymbol{R} = \begin{bmatrix} 0 & 10/3 \\ 0 & 15/2 \end{bmatrix}, \quad \rho(\boldsymbol{B}_{\mathrm{JGS}}) = 7.5 > 1,$$

故 JGS 格式也发散. 当 $\boldsymbol{A} = \boldsymbol{A}_3$ 时,由于

$$\boldsymbol{B}_{\mathrm{J}} = \begin{bmatrix} 0 & -2 & 2 \\ -1 & 0 & -1 \\ -2 & -2 & 0 \end{bmatrix}, \quad \rho(\boldsymbol{B}_{\mathrm{J}}) = 0 < 1,$$

所以 Jacobi 格式收敛;又

$$\boldsymbol{B}_{\mathrm{JGS}} = \begin{bmatrix} 0 & -2 & 2 \\ 0 & 2 & -3 \\ 0 & 0 & 2 \end{bmatrix}, \quad \rho(\boldsymbol{B}_{\mathrm{JGS}}) = 2 > 1,$$

故 JGS 格式发散. 当 $\boldsymbol{A} = \boldsymbol{A}_4$ 时,由于

$$\boldsymbol{B}_{\mathrm{J}} = \begin{bmatrix} 0 & 1/2 & -1/2 \\ -1 & 0 & -1 \\ 1/2 & 1/2 & 0 \end{bmatrix}, \quad \rho(\boldsymbol{B}_{\mathrm{J}}) = \sqrt{1.25} > 1,$$

所以 Jacobi 格式发散;又

$$\boldsymbol{B}_{\mathrm{JGS}} = \begin{bmatrix} 0 & 1/2 & -1/2 \\ 0 & -1/2 & -1/2 \\ 0 & 0 & -1/2 \end{bmatrix}, \quad \rho(\boldsymbol{B}_{\mathrm{JGS}}) = 0.5 < 1,$$

故 JGS 格式收敛.

定理 1.6 设 $\boldsymbol{A}$ 为 Hermite 正定矩阵,且实参数 ω 满足 $0 < \omega < 2$,则 SOR 格式收敛.

推论 1.1 设 $\boldsymbol{A}$ 为 Hermite 正定矩阵,则 JGS 格式收敛.

定理 1.7 设 $\boldsymbol{A}$ 按行(列)严格对角占优,或者弱对角占优且不可约,则当实参数 ω 满足 $0 < \omega \leqslant 1$ 时,SOR 格式收敛.

1.2 变分原理迭代方法

本节约定:矩阵 $\boldsymbol{A} \in \mathbf{R}^{n\times n}$ 对称正定,列向量 $\boldsymbol{b} \in \mathbf{R}^n$.

首先建立求解方程组(1.1)的变分原理. 对于系数矩阵 $\boldsymbol{A}$ 及右端向量 $\boldsymbol{b}$,定义 n 元二次函数为

$$\varphi(\boldsymbol{x}) = \frac{1}{2}[\boldsymbol{x}, \boldsymbol{A}\boldsymbol{x}] - [\boldsymbol{x}, \boldsymbol{b}], \tag{1.11}$$

式中:$\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^{\mathrm{T}} \in \mathbf{R}^n$.

定理 1.8 设 $\boldsymbol{A}$ 为 n 阶实对称正定矩阵,$\boldsymbol{b}$ 为 n 维实列向量,则 $\boldsymbol{x}^* \in \mathbf{R}^n$ 满足 $\boldsymbol{A}\boldsymbol{x}^* = \boldsymbol{b}$ 的充要条件是 $\varphi(\boldsymbol{x}^*) = \min\limits_{\boldsymbol{x}\in\mathbf{R}^n}\varphi(\boldsymbol{x})$.

定理 1.8 将求方程组(1.1)的唯一解问题,转化为求函数 $\varphi(\boldsymbol{x})$ 的唯一极小点问题. 为了找到 $\varphi(\boldsymbol{x})$ 的极小点 $\boldsymbol{x}^*$,可以从任一点 $\boldsymbol{x}^{(k)}$ 出发,沿着某一指定的方向 $\boldsymbol{z}^{(k)} \in \mathbf{R}^n$,搜索下一个近似点 $\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \alpha_k \boldsymbol{z}^{(k)}$,使得 $\varphi(\boldsymbol{x}^{(k+1)})$ 在该方向上达到极小值. 指定 $\boldsymbol{z}^{(k)}$ 的方式不同时,将会得到不同的算法.

令 $\boldsymbol{r}^{(k)} = \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}^{(k)}$,由式(1.11)可得

$$f(\alpha) = \varphi(\boldsymbol{x}^{(k)} + \alpha \boldsymbol{z}^{(k)}) = \varphi(\boldsymbol{x}^{(k)}) - \alpha[\boldsymbol{r}^{(k)}, \boldsymbol{z}^{(k)}] + \frac{\alpha^2}{2}[\boldsymbol{A}\boldsymbol{z}^{(k)}, \boldsymbol{z}^{(k)}],$$

再由 $f'(\alpha) = 0$ 求得 $f(\alpha)$ 的唯一驻点为 $\alpha_k = \dfrac{[\boldsymbol{r}^{(k)}, \boldsymbol{z}^{(k)}]}{[\boldsymbol{A}\boldsymbol{z}^{(k)}, \boldsymbol{z}^{(k)}]}$. 因为

$$f''(\alpha) = [\boldsymbol{A}\boldsymbol{z}^{(k)}, \boldsymbol{z}^{(k)}] > 0 \quad (\boldsymbol{z}^{(k)} \neq \boldsymbol{0}),$$

所以 $f''(\alpha_k) > 0$,故 α_k 是 $f(\alpha)$ 的极小点. 直接计算可得

$$\varphi(\boldsymbol{x}^{(k)} + \alpha_k \boldsymbol{z}^{(k)}) - \varphi(\boldsymbol{x}^{(k)}) = -\frac{1}{2}\frac{[\boldsymbol{r}^{(k)}, \boldsymbol{z}^{(k)}]^2}{[\boldsymbol{A}\boldsymbol{z}^{(k)}, \boldsymbol{z}^{(k)}]}. \tag{1.12}$$

当 $[\boldsymbol{r}^{(k)}, \boldsymbol{z}^{(k)}] \neq 0$,即 $\boldsymbol{z}^{(k)}$ 不与 $\boldsymbol{r}^{(k)}$ 正交时,$\varphi(\boldsymbol{x}^{(k+1)}) < \varphi(\boldsymbol{x}^{(k)})$ 成立.

1.2.1 最速下降法

相对于 $\varphi(\boldsymbol{x}^{(k)})$,由式(1.12)知 $\varphi(\boldsymbol{x}^{(k+1)})$ 的下降量取决于 $\boldsymbol{z}^{(k)}$ 的方向,而与 $\boldsymbol{z}^{(k)}$ 的大小无关. 根据多元微积分知识可知,$\varphi(\boldsymbol{x})$ 在点 $\boldsymbol{x}^{(k)}$ 下降最快的方向是在该点的负梯度方向,即

$$-\operatorname{grad}\varphi(\boldsymbol{x})\big|_{\boldsymbol{x}=\boldsymbol{x}^{(k)}} = \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}^{(k)} = \boldsymbol{r}^{(k)}.$$

取 $\boldsymbol{z}^{(k)} = \boldsymbol{r}^{(k)}$,并构造 $\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \alpha_k \boldsymbol{r}^{(k)}$,可得确定 $\varphi(\boldsymbol{x})$ 的极小点 $\boldsymbol{x}^*$ 的**最速下降法**:任取 $\boldsymbol{x}^{(0)} \in \mathbf{R}^n$,迭代格式为

$$\begin{gathered}\boldsymbol{r}^{(k)} = \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}^{(k)}, \quad \alpha_k = \frac{[\boldsymbol{r}^{(k)}, \boldsymbol{r}^{(k)}]}{[\boldsymbol{A}\boldsymbol{r}^{(k)}, \boldsymbol{r}^{(k)}]}, \\ \boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \alpha_k \boldsymbol{r}^{(k)} \quad (k = 0,1,2,\cdots).\end{gathered} \tag{1.13}$$

定理 1.9 设 $\boldsymbol{A}$ 为 n 阶实对称正定矩阵,λ_1 和 λ_n 分别表示 $\boldsymbol{A}$ 的最大、最小特征值,则由格式(1.13)得到的向量序列 $\{\boldsymbol{x}^{(k)}\}$ 满足

$$\| \boldsymbol{x}^{(k)} - \boldsymbol{x}^* \|_{\boldsymbol{A}} \leqslant \left(\frac{\lambda_1 - \lambda_n}{\lambda_1 + \lambda_n}\right)^k \| \boldsymbol{x}^{(0)} - \boldsymbol{x}^* \|_{\boldsymbol{A}}. \tag{1.14}$$

式(1.14)表明,最速下降法在理论上是收敛的. 但当 $\lambda_1 >> \lambda_n$ 时,$\dfrac{\lambda_1 - \lambda_n}{\lambda_1 + \lambda_n} \approx 1$,格式(1.13)将会收敛缓慢.

1.2.2 共轭梯度法

在最速下降法中,选取 $\boldsymbol{z}^{(k)}$ 为 $\varphi(\boldsymbol{x})$ 在点 $\boldsymbol{x}^{(k)}$ 的最速下降方向,即 $\boldsymbol{z}^{(k)} = \boldsymbol{r}^{(k)}$. 但是,当 $\boldsymbol{r}^{(k)}$ 较小时,由于舍入误差的影响,实际计算得到的 $\boldsymbol{r}^{(k)}$ 会偏离最速下降方向. 这里选取

$$\boldsymbol{z}^{(0)} = \boldsymbol{r}^{(0)}, \quad \boldsymbol{z}^{(k)} = \boldsymbol{r}^{(k)} + \beta_k \boldsymbol{z}^{(k-1)} \quad (k = 1,2,\cdots), \tag{1.15}$$

并要求 $[\boldsymbol{z}^{(k)}, \boldsymbol{A}\boldsymbol{z}^{(k-1)}] = 0$,可求得

$$\beta_k = -\frac{[\boldsymbol{r}^{(k)}, \boldsymbol{A}\boldsymbol{z}^{(k-1)}]}{[\boldsymbol{z}^{(k-1)}, \boldsymbol{A}\boldsymbol{z}^{(k-1)}]}. \tag{1.16}$$

于是可建立求 $\varphi(\boldsymbol{x})$ 的极小点 $\boldsymbol{x}^*$ 的**共轭梯度法**:任取 $\boldsymbol{x}^{(0)} \in \mathbf{R}^n$,迭代格式为

$$\boldsymbol{r}^{(0)} = \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}^{(0)},\quad \boldsymbol{z}^{(0)} = \boldsymbol{r}^{(0)},\quad \alpha_0 = \frac{[\boldsymbol{r}^{(0)},\boldsymbol{z}^{(0)}]}{[\boldsymbol{A}\boldsymbol{z}^{(0)},\boldsymbol{z}^{(0)}]},$$
$$\boldsymbol{x}^{(1)} = \boldsymbol{x}^{(0)} + \alpha_0\boldsymbol{z}^{(0)};$$
$$\boldsymbol{r}^{(k)} = \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}^{(k)},\quad \boldsymbol{z}^{(k)} = \boldsymbol{r}^{(k)} + \beta_k\boldsymbol{z}^{(k-1)},\quad \alpha_k = \frac{[\boldsymbol{r}^{(k)},\boldsymbol{z}^{(k)}]}{[\boldsymbol{A}\boldsymbol{z}^{(k)},\boldsymbol{z}^{(k)}]},$$
$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \alpha_k\boldsymbol{r}^{(k)}\quad (k = 0,1,2,\cdots), \tag{1.17}$$

其中,β_k 由式(1.16)计算. 下面讨论式(1.17)的收敛性[2,6].

定理1.10 设 $\boldsymbol{A}$ 为 n 阶实对称正定矩阵,则由格式(1.17)得到的向量序列 $\{\boldsymbol{r}^{(k)}\}$ 和 $\{\boldsymbol{z}^{(k)}\}$ 满足

$$[\boldsymbol{r}^{(k)},\boldsymbol{z}^{(k-1)}] = 0,\quad [\boldsymbol{r}^{(k)},\boldsymbol{r}^{(l)}] = 0,\quad [\boldsymbol{z}^{(k)},\boldsymbol{A}\boldsymbol{z}^{(l)}] = 0\quad (k \neq l). \tag{1.18}$$

定理1.10表明,如果由格式(1.17)得到的向量 $\boldsymbol{r}^{(0)},\boldsymbol{r}^{(1)},\cdots,\boldsymbol{r}^{(n-1)}$ 都不等于零向量,那么它们线性无关,从而构成 $\mathbf{R}^n$ 的一组基. 再由 $\boldsymbol{r}^{(n)}$ 和它们都正交可得 $\boldsymbol{r}^{(n)} = \mathbf{0}$,即 $\boldsymbol{A}\boldsymbol{x}^{(n)} = \boldsymbol{b}$. 因此,在格式(1.17)中的所有计算没有舍入误差的条件下,最多经过 n 次迭代即可得到精确解 $\boldsymbol{x}^*$. 但是,由于计算过程不可避免的舍入误差,使用格式(1.17)并不能在 n 次迭代后给出 $\boldsymbol{x}^*$,所以只能把它当作迭代方法使用.

根据定理1.10,有

$$[\boldsymbol{r}^{(k)},\boldsymbol{z}^{(k)}] = [\boldsymbol{r}^{(k)},\boldsymbol{r}^{(k)} + \beta_k\boldsymbol{z}^{(k-1)}] = [\boldsymbol{r}^{(k)},\boldsymbol{r}^{(k)}].$$

因此,可将迭代格式(1.17)中的 α_k、β_k 改写为

$$\alpha_k = \frac{[\boldsymbol{r}^{(k)},\boldsymbol{z}^{(k)}]}{[\boldsymbol{A}\boldsymbol{z}^{(k)},\boldsymbol{z}^{(k)}]} = \frac{[\boldsymbol{r}^{(k)},\boldsymbol{r}^{(k)}]}{[\boldsymbol{A}\boldsymbol{z}^{(k)},\boldsymbol{z}^{(k)}]},$$
$$\beta_k = -\frac{[\boldsymbol{r}^{(k)},\boldsymbol{A}\boldsymbol{z}^{(k-1)}]}{[\boldsymbol{z}^{(k-1)},\boldsymbol{A}\boldsymbol{z}^{(k-1)}]} = -\frac{[\boldsymbol{r}^{(k)},\boldsymbol{r}^{(k-1)} - \boldsymbol{r}^{(k)}]}{[\boldsymbol{z}^{(k-1)},\boldsymbol{r}^{(k-1)} - \boldsymbol{r}^{(k)}]} = \frac{[\boldsymbol{r}^{(k)},\boldsymbol{r}^{(k)}]}{[\boldsymbol{z}^{(k-1)},\boldsymbol{r}^{(k-1)}]} = \frac{[\boldsymbol{r}^{(k)},\boldsymbol{r}^{(k)}]}{[\boldsymbol{r}^{(k-1)},\boldsymbol{r}^{(k-1)}]}.$$

定理1.11 设 $\boldsymbol{A}$ 为 n 阶实对称正定矩阵,λ_1 和 λ_n 分别表示 $\boldsymbol{A}$ 的最大、最小特征值,则由格式(1.17)得到的向量序列 $\{\boldsymbol{x}^{(k)}\}$ 满足

$$\|\boldsymbol{x}^{(k)} - \boldsymbol{x}^*\|_A \leqslant 2\left(\frac{\sqrt{\lambda_1} - \sqrt{\lambda_n}}{\sqrt{\lambda_1} + \sqrt{\lambda_n}}\right)^k \|\boldsymbol{x}^{(0)} - \boldsymbol{x}^*\|_A. \tag{1.19}$$

比较式(1.19)和式(1.14),由于

$$\left(\frac{\lambda_1 - \lambda_n}{\lambda_1 + \lambda_n}\right)^k = \left(\frac{\sqrt{\lambda_1} - \sqrt{\lambda_n}}{\sqrt{\lambda_1} + \sqrt{\lambda_n}}\right)^k \left(\frac{(\sqrt{\lambda_1} + \sqrt{\lambda_n})^2}{\lambda_1 + \lambda_n}\right)^k$$
$$= 2\left(\frac{\sqrt{\lambda_1} - \sqrt{\lambda_n}}{\sqrt{\lambda_1} + \sqrt{\lambda_n}}\right)^k \left(\frac{1}{2}\left(1 + \frac{2\sqrt{\lambda_1\lambda_n}}{\lambda_1 + \lambda_n}\right)^k\right),$$

所以共轭梯度法比最速下降法的收敛性好得多.

1.3 整体校正加速方法

使用迭代方法求线性代数方程组,或者线性矩阵方程的唯一解时,迭代过程可能收敛,也可能不收敛. 即便收敛,也可能收敛得很慢. 无论是哪种情况,都希望找到一种改进