多元非线性方程组迭代解法

J. M. 奥特加　W. C. 莱因博尔特　著

科 学 出 版 社

内 容 简 介

本书系统地介绍n阶非线性方程组的基本理论成果，并且对求这类方程数值解的几种主要迭代方法进行了分析. 全书共分五部分，即本书所需要的基础知识，非结构性存在定理，迭代法，局部收敛性，半局部收敛性和整体收敛性等. 书中对所有结论都给出了详细的证明，并列有大量习题，每节末还有注记.

本书可供从事计算数学和计算机应用的科技工作者、高等院校有关专业的教师、研究生以及高年级学生参考.

J. M. Ortega and W. C. Rheinboldt

ITERATIVE SOLUTION OF NONLINEAR EQUATIONS IN SEVERAL VARIABLES

ACADEMIC PRESS, 1970

多元非线性方程组迭代解法

J. M. 奥特加 W. C. 莱因博尔特 著

朱季纳 译

孙念增 进 友 校

责任编辑 向安全

科学出版社 出版

北京朝阳门内大街 137 号

中国科学院印刷厂印刷

新华书店北京发行所发行 各地新华书店经售

*

1983年10月第一版 开本：850×1168 1/32

1983年10月第一次印刷 印张：19

印数：0001—8,500 字数：495,000

统一书号：13031・2387

本社书号：3265・13—1

定价： 3.50 元

序　　言

本书是根据我们和几位博士学位研究生在过去五年中所从事的研究工作编写而成的；大部分内容也包含在这一时期我们在 Maryland 大学开的研究生课程中．我们的目的是对 n 阶非线性方程组的基本理论结果作一个概括性的叙述，同时对它们的数值解的主要迭代方法进行分析．对这个飞速发展的领域作这样一个详细介绍，看来是需要的，不仅对这个领域内的工作者有益，而且对特殊结果或方法有兴趣或需要的读者也有帮助．同时我们也希望在这里为这个领域内的研究生提供一本数值分析课程的教材，为了达到这个目的，我们力图使本书内容尽可能地自成体系，充分详细地证明所有结果，并在全书中安排了许多习题．为了使本书也可用于查阅文献，我们在每节末补充了一批“注记”，指出所引用的文献，讨论其它有关的结果，并指出正文中结果的各种可能推广．此外，书末还附有这方面的详细文献目录．

本书假定读者已具备相当于标准的高等多元微积分和线性代数课程水平的预备知识．在第二和第三章中按本书所需要的形式复习和收集了其中一些材料．熟悉一些解一个未知数的单个方程和线性方程组的基本方法，也是有帮助的，但不是必不可少的．特别是在第二章，收集了线性方程组数值解法理论中我们所需要的全部基础知识．

根据需要，我们必须作一些规定以限制本书的内容．我们没有专门讲述一阶方程或线性方程组的迭代解法，因为前一部分内容已包括在 Ostrowski [1966] 和 Traub [1964] 的优秀专著中，后一部分内容已包括在 Forsythe 和 Wasow [1960] 以及 Varga [1962] 的优秀著作中．除少数地方外，我们限于求有 n 个实未知数的 n 个实方程的孤立解(没有其它约束)的问题．“求解”有 n 个未知数

的 m 个方程的问题，当 $m > n$ 时是逼近论的一个课题，而当 $m < n$ 时看来没有什么独立的意义. 引进一组附加条件以约束所要求的解，需要有不同的方法，在一些情况下，它是非线性规划的问题，而对它的一般形式还很少了解. 类似地，非孤立解问题几乎是完全尚未研究的领域. 我们不讨论在计算公式中需要二阶或更高阶导数的迭代法，因为对这类方法的分析搞不出什么名堂，而且十分烦琐，更重要的是因为计算一个将 R^n 映入它自身的映射的 k 阶导数一般需要算出 n^{k+1} 次函数值. 因而我们认为除特殊问题外，需要高于一阶导数的数值计算方法是不可取的. 最后，我们没有安排数值计算的例子. 如果将各种方法用于计算那些理论上已保证不存在任何困难的偶然的例题上，看来收获会很小. 为了对一种迭代过程的数值特性有一个深入的了解，需要进行广泛的数值计算，不仅要系统地改变所解的方程和它的阶数，而且对每一种情形还要取很多不同的初始近似. 据我们所知，至今尚未充分进行过这类计算工作，而我们自己的计算实践表明，存在的问题比能解决的还多. 改变方程、阶数和初始数据对计算后果的影响，在实践和理论两方面都仍然知道得很少，特别是，对各类计算误差的影响，只取得了很少的结果.

本书内容的主要轮廓也许是由于这样一个规定造成的，即全书只限于有限维情形，而不考虑在这个领域内多数研究结果通常是在更一般情况下给出的. 不过，我们尽可能多地介绍一些近代研究的结果，使得推广到 Banach 空间上的算子立即变得很清楚. 我们相信这样做有一大优点，可使本书对具有广泛的泛函分析基础或完全没有这种基础的读者都是有意义的，而且是可接受的. 此外，在假定读者熟悉必要的术语的条件下，在"注记"中指出了到无穷维空间的推广. 关于某些这类专题在泛函分析背景下的论述，读者可参阅下列各书： Collatz [1964], Goldstein [1967], Kantorovich 和 Akilov [1959], Rall [1969] 以及Vainberg [1956].

符　号　表

空间(见 2.1 节)

R^n　实 n 维空间

C^n　复 n 维空间

$L(R^n, R^m)$, $L(R^n)$　从 R^n 到 R^m 或从 R^n 到 R^n 的线性算子的线性空间

$R^n \times R^m$　乘积空间

$(R^n)^m$　m 重乘积空间 $R^n \times \cdots \times R^n$

向量和集合(见第二章)

$e^1, \cdots, e^n$　R^n 的坐标向量

$x = (x_1, \cdots, x_n)^T$　分量为 x_i 的列向量

x^T　x 的转置

$\{x^1, \cdots, x^m\}$　m 个向量的集合

$\{x^k\}$　向量的序列

$\|\cdot\|$　R^n 上的任一种范数

$\|\cdot\|_p$　R^n 上的 l_p-范数，$1 \leqslant p \leqslant \infty$

(x, y)　R^n 上的一种内积

$S(x, r)$　开球 $\{y \in R^n \mid \|y - x\| < r\}$

$\bar{S}(x, r)$　闭球 $\{y \in R^n \mid \|y - x\| \leqslant r\}$

$\bar{S}$　集合 S 的闭包

$\dot{S}$　集合 S 的边界

$\mathrm{int}(S)$　集合 S 的内部

$[x, y]$　集合 $\{z \in R^n \mid z = tx + (1 - t)y, t \in [0, 1]\}$

$x \leqslant y$　偏序 $x_i \leqslant y_i, i = 1, \cdots, n$

$\langle x, y \rangle$　集合 $\{z \in R^n \mid x \leqslant z \leqslant y\}$

$|x|$　分量为 $|x_i|, i = 1, \cdots, n$ 的向量

矩阵　(见第二章)

$A = (a_{ij})$　元素为 a_{ij} 的 $n \times n$ 矩阵

A^{-1}　A 的逆矩阵

A^{T}　A 的转置

$\det A$　A 的行列式

A^{+}　A 的一个广义逆矩阵

A^{k}　A 的 k 次幂

$\{A_k\}$　矩阵的序列

$\rho(A)$　A 的谱半径

$\|A\|$　A 的任一种范数

$\|A\|_p$　A 的 l_p-范数，$1\leqslant p\leqslant\infty$

$\operatorname{rank} A$　A 的秩

I　单位矩阵

$A\leqslant B$　偏序 $a_{ij}\leqslant b_{ij}$, $i, j=1,\cdots,n$

$(a^1,\cdots,a^n)$　以 $a^1,\cdots,a^n$ 为列的矩阵

$\operatorname{diag}(a_1,\cdots,a_n)$　元素为 $a_1,\cdots,a_n$ 的对角矩阵

函数和导数　(见第三章)

$F: D\subset R^n\to R^m$　定义域 D 在 R^n 内值域在 R^m 内的映射

$F(D)$ 或 FD　集合 $\{y\in R^m \mid y=Fx,\ x\in D\}$

$F'(x)$, $F''(x)$　F 在 x 的一阶或二阶 (G 或 F) 导数

F^p　F 的 p 次幂

F^{-1}　F 的逆

F_U　F 对集合 U 的限制

$\partial_i f$　对第 i 个变量的偏导数

$\partial_i F$　向量偏导数(见 5.2.2 节)

$F(\cdot, y)$　取定 y 后的映射

$H_g(x)$　g 在 x 的 Hesse 矩阵

Δ　Laplace 算子

其它

$\in$　元素属于

$\subset$　集合含于

$\cup$, $\cap$　集合的并或交

$S_1\sim S_2$　集合 S_1 和 S_2 的差

$\forall$　对所有的

Π, Σ　积，和

δ_{ij}　Kronecker 符号

$\operatorname{sgn} x$　函数 $\operatorname{sgn} x = 1$（当 $x \geqslant 0$ 时），$\operatorname{sgn} x = -1$（当 $x < 0$ 时）

$\lim\limits_{t \to a^-}$　取 $t < a$，$t \longrightarrow a$ 的单侧极限

$\mathscr{I}$　一个一般的迭代过程

$C(\mathscr{I}, x^*)$　由$\mathscr{I}$生成的收敛于 x^* 的序列的集合

$$\left.\begin{array}{l} R_p\{x^k\},\ R_p(\mathscr{I}, x^*) \\ Q_p\{x^k\},\ Q_p(\mathscr{I}, x^*) \\ O_R\{x^k\},\ O_R(\mathscr{I}, x^*) \\ O_Q\{x^k\},\ O_Q(\mathscr{I}, x^*) \end{array}\right\}$$ 收敛速度的记号(见第九章)

目　录

第二部分 非构造性的存在定理

第三部分 迭 代 法

第四部分 局部收敛性

第五部分 半局部收敛性和整体收敛性

引　　言

我们用 R^n 表示由所有列向量

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

组成的实 n 维线性空间. 不过，除了一些与坐标有关的迭代过程(如 7.4 节中介绍的 SOR 法)外，所有的讨论实质上都与基无关，因而 R^n 也可看作抽象的实 n 维线性空间. 我们将用带或不带上标的小写拉丁字母表示 R^n 中的向量，而用下标表示这些向量的分量；例如，x_i^k 是向量 x^k 的第 i 个分量. 对有关向量和矩阵记号的更详细规定，将在第二章中给出.

通常假定已赋予 R^n 某种无需特别说明的范数(见 2.2 节)，只有当结果和一种特殊的范数有关时，我们才明确说明这个范数. 假定读者已熟悉 R^n 上基本的拓扑概念，如开集、闭集、紧集、一点的邻域、极限和 Cauchy 序列，函数的连续性和一致连续性等. 如果读者只对普通的 Euclid 范数熟悉这些概念，则重要的是应注意到，利用范数等价定理 2.2.1 (由它可得 R^n 上所有拓扑的讨论都与范数无关)，可以将这些概念直接移植到任一种范数.

集合 $S \subset R^n$ 的闭包、边界和内部分别记作 $\bar{S}$, $\dot{S}$ 和 $\mathrm{int}(S)$. R^n 的特别重要的子集是以 x^0 为中心，$r > 0$ 为半径的开球和闭球(对某种范数 $\|\cdot\|$)：

$$S(x^0, r) = \{x \in R^n \mid \|x - x^0\| < r\},$$

$$\bar{S}(x^0, r) = \{x \in R^n \mid \|x - x^0\| \leqslant r\}.$$

定义域 D 在 R^n 内，值域在 R^m 内的函数 F (也叫映射或算子)将记作 $F: D \subset R^n \to R^m$，如果集合 D 和 Q 的维数是明显的，有时也记作 $F: D \to Q$. 当 $m > 1$ 时，$F: D \subset R^n \to R^m$ 的分量记作

$f_1, \cdots, f_m$，并用下面列向量表示 $Fx \in R^m$：

$$Fx = \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix}.$$

利用这些记号规定，我们讨论的问题是求解方程组

$$f_i(x_1, \cdots, x_n) = y_i, \quad i = 1, \cdots, n, \tag{1}$$

或缩写成

$$Fx = y, \tag{2}$$

其中 $F: D \subset R^n \to R^n$ 是给定的算子，$y \in R^n$ 是一个取定的向量．将向量 y 吸收到 F 中通常没有什么限制，故可只考虑方程

$$Fx = 0. \tag{3}$$

在着手求 (3) 的解之前，重要的是要认识到，这样一个问题甚至可能没有解，或者可能有任意多个解．为了说明这一点，考虑方程组

$$f_1(x_1, x_2) = 0, \quad f_2(x_1, x_2) = 0. \tag{4}$$

这两个方程中每一个表示 R^2 中一条(不一定连续)曲线，所以 (4) 的解是这两条曲线的交点．如果在简单的例

$$f_1(x_1, x_2) \equiv x_1^2 - x_2 + \alpha, \quad f_2(x_1, x_2) \equiv -x_1 + x_2^2 + \alpha$$

中，实参数 α 在 $+1$ 和 -1 之间变化，就有下列情形(见图 I.1)：

(a) $\alpha = 1$： 无解．

(b) $\alpha = \frac{1}{4}$： 一个解，$x_1 = x_2 = \frac{1}{2}$．

(c) $\alpha = 0$： 两个解，$x_1 = x_2 = 0$；$x_1 = x_2 = 1$．

(d) $\alpha = -1$：四个解，$x_1 = -1$，$x_2 = 0$；$x_1 = 0$，$x_2 = -1$；

$$x_1 = x_2 = \frac{1}{2}(1 \pm \sqrt{5}).$$

类似地，方程组

$$f_1(x_1, x_2) \equiv \frac{1}{2} x_1 \left[\sin\left(\frac{1}{2}\pi x_1\right)\right] - x_2 = 0,$$

$$f_2(x_1, x_2) \equiv x_2^2 - x_1 + 1 = 0$$