# 关 肇 直 文 集

《关肇直文集》编辑小组 编

科 学 出 版 社

1986

# 内 容 简 介

关肇直教授是我国已故著名数学家．他的研究领域包括泛函分析、控制理论、数学物理和数学史等方面．本文集收集了他的二十五篇论文．

本书读者对象是从事数学研究的人员，以及大学高年级学生、研究生、教师及有关的工程技术人员．


**关 肇 直 文 集**

《关肇直文集》编辑小组 编

责任编辑 吕 虹 张鸿林

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

*

1986年4月第一版 开本：787×1092 1/16

1986年4月第一次印刷 印张：11 3/4

印数：精 1—1,500 插页：精 3 平 1

平 1—1,500 字数：270,000

统一书号：13031·3101

本社书号：4420·13—1

定价：布脊精装 3.85 元

平 装 2.85 元

关 肇 直

(1919—1982)

# 序　言

关肇直同志，广东省南海县人，生于1919年2月13日，1941年毕业于燕京大学数学系．那是抗日战争年代，中国知识分子颠沛流离，肇直同志亦不得不离开当时的北平而负笈华西．在燕京的年月，是他开始治学和从事革命活动的年月．

肇直同志的情趣、学识和工作是多方面的，但核心是发展中国的数学．他的第一篇论文发表于 The Mathematical Gazette (1939)，题为“Chinese Unit of Length in the Early Ching Dynasty”．他的哲学、历史和文学的素养都极好，因而在此后四十年内，他仍不时回到中外数学史这一研究领域中来．

1947年他赴法留学，1949年新中国诞生，“漫卷诗书喜欲狂”，他欣然束装归国．在法时间虽不长，但因受 Fréchet 等人的熏陶，使得泛函分析成为他终身致力的数学分支之一．他曾从事数值分析、控制理论和数学物理等方面的工作，但从其中都可以见到他对泛函分析的爱好和造诣．

1956年他在“解非线性函数方程的最速下降法”一文中，研究了无穷维空间中非线性方程的近似解法，证明了 Hilbert 空间中非线性方程的最速下降法的收敛性．在这篇论文中，他实际上引进了“单调算子”的概念．二十余年来，国际上在这方面所引起的兴趣和取得的成就，足以说明单调算子是一类重要的非线性算子，于此可见肇直同志洞察力之强．他很珍惜他的这一发现．1980年他积劳成疾，在病榻上，他说：“如果不是为了其它工作的需要，我会对单调算子做更多的工作”．

由于他为发展国家科学事业的责任感，以及他的秉性和远见卓识，他甘于开疆拓土而不安于一域一邑的治理．他一贯强调理论与实际联系，并指出正因为与实际联系，才更需要加强理论的研究，他很早就关心跨学科工作的发展．

五十年代中国正在开展原子能事业，同时物理学中出现了新发现——激光，这些都是肇直同志关心的问题，他研究了气体激光中非零本征值的存在性．对原子能科学中的中子迁移问题，他给予了很大注意．当时研究这个问题的一个重要方法——Case 方法还缺乏理论基础，引起了较多理论工作者的注意，肇直同志首先给予了严格的论证．1963年他完成了题为“关于中子迁移理论中出现的一类本征值问题”的论文．1973年在国际上才有类似的工作出现，我们不能不惋惜肇直同志的这一领先的工作，推迟了十多年，方能首次在这里发表．

从六十年代开始，特别是从1966年以后，肇直同志的精力和时间，绝大部分花费在开展中国的控制理论研究．这个时期是近代控制理论开创和发展的时期，线性系统的工作很快就臻趋成熟，由于肇直同志在泛函分析方面的造诣，自然会使他关心分布参数系统和非线性系统．他对弹性振动的控制系统，提出用线性算子谱的紧扰动的方法去研究，取得了很好的结果．他还用泛函分析等数学工具，去研究工程实际中所提出的量化误差、谐波平衡法等理论问题，并对许多种导航、制导、控制系统提出理论分析和原理方案．这些结果，只有少数见诸于公开刊物．

从控制理论到系统理论是国际许多知名学者，也是肇直同志学术思想发展的途径．在"复杂系统的辨识与控制提纲"一文中，他提到了 Prigogine 有关非平衡热力学的工作以及 Thom 的突变理论，在他看来这是系统科学的主要内容．1982 年，当他病情已经相当恶化时，他还表示，要等身体恢复健康之后，着意致力于这方面的研究．1982 年 11 月 12 日，肇直同志与世长逝，这项也许将是他一生中最重要的工作，刚刚开始，即宣告结束．

这本论文集选择了肇直同志的论文共二十五篇，这是他的文章中的一部分．事实上，肇直同志以更大的手笔，谱写了更为重要的文章，这就是他倡导的学风，他身体力行的主张，他参与制定的规划以及由他指导或受他影响的一些后学的工作，当然，这些都无法以文字的形式载入本集．但是，它们对中国的数学界产生了深远的影响．

吴文俊　许国志

1984 年 2 月

# 目　　录

# 清初中国的长度单位*

## 前　　言

北京近郊燕京大学北面不远，是著名的圆明园的遗址．圆明园是二百多年前清朝康熙皇帝统治期间修造的宫殿．康熙后来赐给了儿子雍正，雍正又做了修整和扩建．修造时中国的长度单位也许同现代的长度单位不同，因而用数学方法确定古时的长度单位在数学和考古学上都是使人感兴趣的．另外，还有一桩使人感兴趣的事情：残存的雕花石块原是到中国传教的耶稣会传教士设计的西式宫殿的一部分．他们设计时用的是中国长度单位制还是法国长度单位制呢？

## 方　　法

本研究所用的方法同 G. F. Cramer[1] 的方法相似．

如果取一组测量值，这组值不是任何单位长度的倍数，再沿一根轴标出这组值，所得的点就不会聚集：我们会发现它们的分布是随机的．若任取数 $N$，将其倍数标在同一根轴上，再在第二组点中每点的两侧各取长度为 $n(2n < N)$的小区间，那么第一组点既可能在双区间 $2n$ 的里面，也可能在外面．所以，任意一组点中大约有 $2n/N$ 个位于区间 $2n$ 之内．由此，我们定义预测分式为 $p = 2n/N$．

另一方面，如果这组测量值是某个单位长度的近似倍数，这些点就会聚集．此时，若 $N$ 的值能够取得接近于这个单位长度的某一倍数，区间 $2n$ 内的点就会多得多．设 $M$ 为点的总数，$m$ 为区间 $2n$ 内点的个数，我们定义实际分式为 $P = m/M$，定义两式之比为 $R = P/p$．

若 $N$ 近似于单位长度的某一倍数，就有 $P > p$ 和 $R > 1$．反之，若 $N$ 的值与单位长度相差较大，就有 $P < p$ 和 $R < 1$．$R$ 的值非常大就表明 $N$ 与单位长度相差无几．

同样地，若将 $N$ 取为 $2N$，$3N$ 等等，假使 $N$ 接近单位长度，$R$ 的值亦会很大．

## 步　　骤

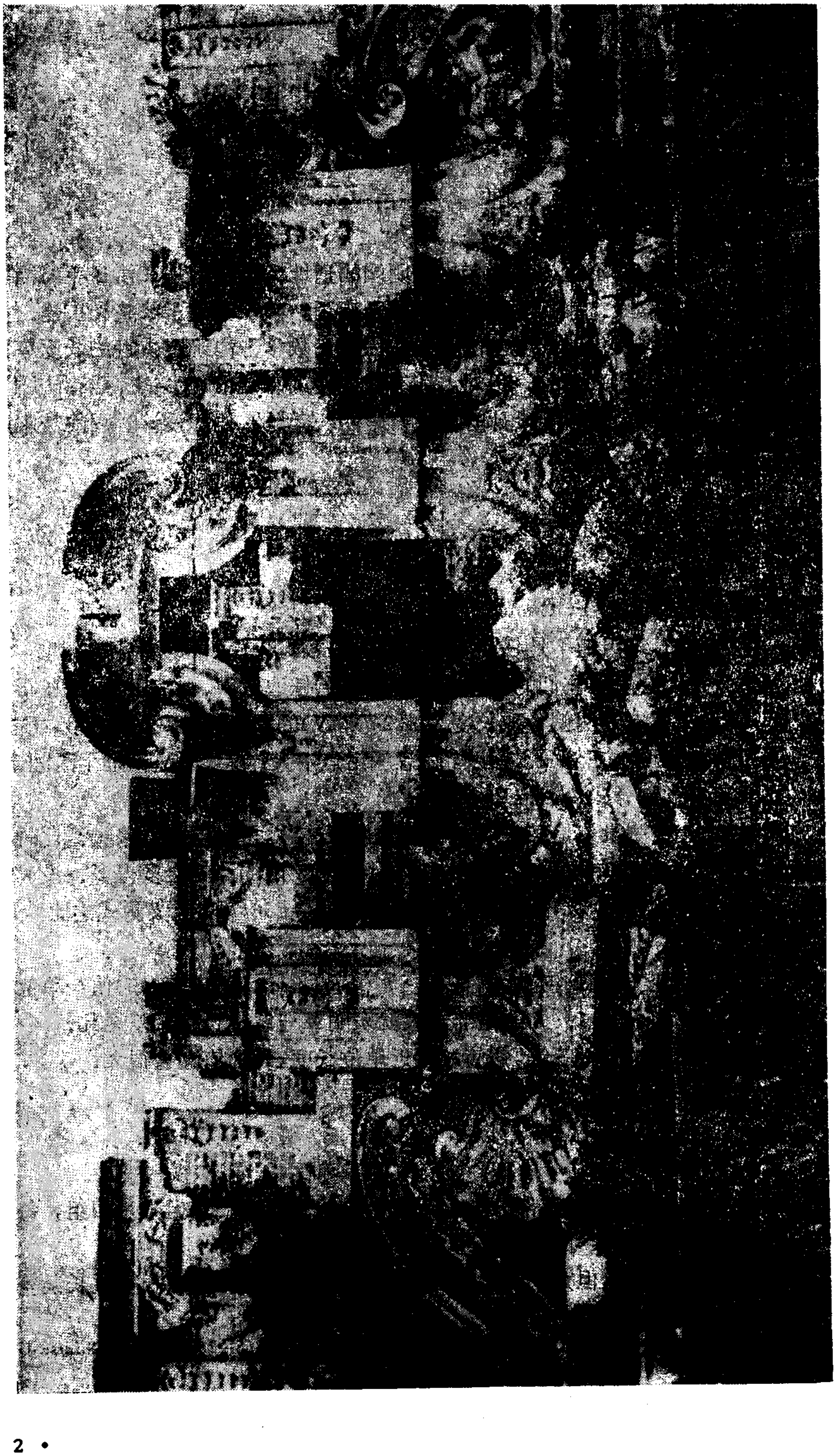
下述工作是由燕京大学数学系部分学生最近完成的．将八个学生分为四组，每组在指定区域内测量掉落石块的尺寸和装饰．

总共测了大约三百个值，但太小的尺寸当时的工匠可能不去量，而长度过长误差又可

---

* 原载：The Mathematical Gazette, Vol. XXIII, No. 225, July, 1939, 268—269，题为："Chinese Unit of Length in the Early Ch'ing Dynasty"．

1) "Determination of a Mayan Unit of Linear Measurement", Amer. Math. Monthly, June, 1938.



圆明园部分遗迹，测量是在此地进行的

圆明园该地原貌

能比较大，所以小于 5 厘米和大于 150 厘米的值都剔去了，测量值的总数因此减为 233 个。复查表明，我们的测量误差肯定小于 0.5 厘米。

## 结　　果

我们取 $n = 0.6$ 厘米。

从图上看，毗邻聚点间的平均距离约为 3 厘米。因此，首先试取 $N$ 在 3 厘米附近的值。我们使用了描图纸以减轻确定区间和数点的工作。所得结果如下：

| $N$(厘米) | $m$ | $R$ |
|---|---|---|
| 3.0 | 102 | 1.09 |
| 3.1 | 96 | 1.07 |
| 3.2 | 130 | 1.49 |
| 3.3 | 103 | 1.22 |
| 3.4 | 73 | 0.89 |
| 1.6 | 187 | 1.07 |
| 2.4 | 113 | 0.97 |
| 6.4 | 90 | 2.06 |
| 9.6 | 46 | 1.58 |

## 讨　　论

从表上可以看到，$N = 3.2, 6.4, 9.6$ 时 $R$ 的值都特别大。$N = 1.6$ 时 $R = 1.07$，因此 1.6 不可能是单位长度。$N$ 取 3.2 附近的值，即 3.0，3.1，3.3 时，$R$ 的值也比较大。这一点应该预料到，如果我们考虑到工匠原有的误差及风化的影响。随意对 $N$ 取其它的值，如 2.4 和 3.4，则 $R < 1$。因此我们猜想 3.2 厘米是古时的单位长度。同 6.4 或 9.6 比较，我们选取 3.2，那是由于下述原因。

现在使用的标准中国尺恰是一米的三分之一，因此中国寸(一尺的十分之一)大约是 3.3 厘米。但这只是近年的事情。1907 年，即民国前四年，满清政府采用 32 厘米作为标准市尺[1]。这一标准取自康熙四十三年(1704 年)造的一个铁瓮底座的尺寸，与当时文件记录的尺寸是一致的。

这个单位长度与我们得到的结果完全一样。64 厘米处是一密集点表明选取 3.2 的误差可能小于 0.1 厘米。

所以我们可以得出结论，尽管设计是法国式样，但却采用了中国长度单位制。这一长度单位制从那时到最近政府对其调整为止，没有多大改变。

（刘峰　译　李文林　校）

1) History of Chinese System of Weights and Measures, by Wu Cheng Lo.

# 关于解函数方程的牛顿方法的一点注记*

本文把 L. Collatz[1] 就复数域场合所论的关于解代数与超越方程的“简化”牛顿法推广到一般 Banach 空间的情形;并且由此,修正并较简单地推导出 I. Fenyö[2] 的两个结果以及 M. Stein[7] 的一个结果.

设 $F(x)$ 是由 Banach 空间 $X$ 到 Banach 空间 $Y$ 中的任意运算子. 我们将设在所论的 $X$ 中某区域 $G$ 内 $F(x)$ 具有连续的 Fréchet 导式 $F'(x)$[4],[5]. 我们求解方程

$$F(x) = 0. \tag{1}$$

设 $F_1(x)$ 对于每个 $x \in G$ 是由 $Y$ 到 $X$ 中的有逆线性有界运算子. 我们令

$$K(x) = x + F_1(x)F(x). \tag{2}$$

于是(1)等价于方程

$$x = K(x). \tag{3}$$

假定 $K$ 在 $G$ 中满足 Lipschitz 条件: 即存在常数 $\alpha, 0 < \alpha < 1$, 使

$$\|K(x_1) - K(x_2)\| \leqslant \alpha\|x_1 - x_2\|, \qquad x_1, x_2 \in G. \tag{4}$$

又设

$$x_{k+1} = K(x_k), \tag{5}$$

即

$$x_{k+1} = x_k + F_1(x_k)F(x_k). \tag{6}$$

设球 $S$

$$\|x - x_1\| \leqslant \frac{\alpha}{1-\alpha}\|x_1 - x_0\| \tag{7}$$

完全含于区域 $G$ 中.

容易看出,如果已证得 $x, x_1, \cdots, x_{k-1}, x_k \in S$, 则依(4),(5)

$$\|x_{k+1} - x_k\| \leqslant \alpha\|x_k - x_{k-1}\| \leqslant \alpha^k\|x_1 - x_0\|,$$

$$\|x_{k+1} - x_1\| \leqslant \|x_{k+1} - x_k\| + \|x_k - x_{k-1}\| + \cdots + \|x_2 - x_1\|$$

$$\leqslant (\alpha^k + \cdots + \alpha)\|x_1 - x_0\| < \frac{\alpha}{1-\alpha}\|x_1 - x_0\|,$$

即 $x_{k+1}$ 也 $\in S$, 依归纳法看出一切 $x_k \in S$, 且

$$\|x_{k+p} - x_k\| \leqslant \alpha^k(1 + \alpha + \cdots + \alpha^{p-1}) \cdot \|x_1 - x_0\| < \frac{\alpha^k}{1-\alpha}\|x_1 - x_0\|,$$

从而 $x_k$ 收敛于一元 $x^*$, 而 $x^*$ 也含于球 $S$ 中. 设 $F_1(x)$ 与 $F(x)$ 都依 $x$ 连续, 则由(5)取极限即得

$$x^* = K(x^*),$$

---

* 原载: 数学进展, **2**:2(1956), 290—295.

从而 $x^*$ 即所求之解.

在实用上,须选一合适的 $F_1(x)$. 我们考察两种情形,各相应于 Канторович[3],[4] 所谓牛顿程序及修改的牛顿程序.

**1.** 设 $F'(x)$存在且连续,而 $F'(x)^{-1}$ 在球 $\|x-x_0\| \leqslant r$ 中存在1),并且设在这个球(今后表示成 $S(x_0;r)$)中,

$$\|F'(x_1)-F'(x_2)\| \leqslant C\|x_1-x_2\|, \qquad x_1,x_2 \in S(x_0;r). \tag{8}$$

又设

$$\|F'(x_0)^{-1}\| \leqslant B, \tag{9}$$

且设

$$3BCr < 1. \tag{10}$$

由于$F(x)$的连续性,可以设初始近似 $x_0$ 足够好,且取 $r$ 足够小,使在 $S(x_0;r)$ 中,

$$\|F(x)\| \leqslant \eta \equiv \frac{2-3BCr}{11B}r. \tag{11}$$

对于 $x_1,x_2 \in S(x_0;r)$ 依(8)及 [5], 305 页公式(9)与(9')[汉译本 331—332 页],

$$\|F(x_2)-F(x_1)-F'(x_1)(x_1-x_2)\| \leqslant \int_0^1 \|F'(x_1+t(x_2-x_1))$$
$$-F'(x_1)\|dt\|x_1-x_2\| \leqslant C\cdot\|x_1-x_2\|^2\int_0^1 t\,dt \leqslant \frac{1}{2}C\|x_1-x_2\|^2. \tag{12}$$

又

$$\|F'(x_1)^{-1}-F'(x_2)^{-1}\| = \|F'(x_1)^{-1}\|\cdot\|F'(x_2)-F'(x_1)\|\cdot\|F'(x_2)^{-1}\|$$
$$\leqslant C\frac{9}{4}B^2\|x_2-x_1\|, \tag{13}$$

因为依照 Fenyö[2] 的证明,对于 $x \in S(x_0;r)$,

$$\|F'(x)^{-1}\| \leqslant \frac{3}{2}B. \tag{14}$$

于是对于 $x_1,x_2 \in S(x_0;r)$, 如果令

$$F_1(x) = -F'(x)^{-1}, \tag{15}$$

即在(2)中取

$$K(x) = x - F'(x)^{-1}F(x),$$

则依(11),(12),(13),(14),对于 $x_1,x_2 \in S(x_0;r)$,

$$\begin{aligned}\|K(x_1)-K(x_2)\| &= \|x_1-F'(x_1)^{-1}F(x_1)-x_2+F'(x_2)^{-1}F(x_2)\| \\ &\leqslant \|x_1-F'(x_1)^{-1}[F(x_1)-F(x_2)]-x_2 \\ &\quad -[F'(x_1)^{-1}-F'(x_2)^{-1}]F(x_2)\| \\ &\leqslant \|F'(x_1)^{-1}[F(x_2)-F(x_1)-F'(x_1)(x_2-x_1)]\| \\ &\quad +\|F'(x_1)^{-1}-F'(x_2)^{-1}\|\cdot\|F(x_2)\| \\ &\leqslant \frac{3}{4}BC\|x_1-x_2\|^2+\frac{9}{4}B^2C\eta\|x_1-x_2\|\end{aligned}$$

---

1) 依照 Fenyö[2], 只须设 $F'(x_0)^{-1}$ 存在,而于是 $F'(x)^{-1}$ 在全球 $S(x_0;r)$ 中的存在可以由此及(8),(9),(10)推出.

$$\leqslant \frac{3}{2} BC \left[ r + \frac{3}{2} B\eta \right] \cdot \|x_1 - x_2\|,$$

即依(11)及(10),此时

$$\alpha = \frac{3}{2} BC \left[ r + \frac{3}{2} B\eta \right] = \frac{3}{2} BC \left[ r + \frac{3[2 - 3BCr]}{22} r \right]$$

$$\leqslant \frac{1}{2} \left[ 1 + \frac{6}{22} \right] = \frac{7}{11} < 1.$$

又

$$\|x_1 - x\| \leqslant \frac{\alpha}{1-\alpha} \|x_1 - x_0\| \Rightarrow \|x - x_0\|$$

$$\leqslant \frac{\alpha}{1-\alpha} \|x_1 - x_0\| + \|x_1 - x_0\| = \frac{1}{1-\alpha} \|x_1 - x_0\|,$$

而因

$$\|x_1 - x_0\| \leqslant \|F'(x_0)^{-1}\| \cdot \|F(x_0)\| \leqslant B\eta,$$

所以

$$\frac{1}{1-\alpha} \|x_1 - x_0\| \leqslant \frac{B\eta}{1-\alpha} \leqslant \frac{B\eta}{1 - \frac{7}{11}} = \frac{11B\eta}{4} < \frac{2 - 3BCr}{4} r < r,$$

即在这情形下,球(7)确完全含于区域 $G \equiv S(x_0; r)$ 中. 于是依上述的一般结果,牛顿程序在 Fenyö 所考虑的情形([2]中的定理 1)之下收敛于方程(1)的解[1)].

**2.** 关于修改的牛顿过程的问题,更简单得多. 这时我们取

$$F_1(x) = -A^{-1}, \tag{16}$$

其中

$$A = F'(x_0). \tag{17}$$

依照 Fenyö, 我们假定在 $S(x_0;r)$ 中,

$$\|F'(x_0)^{-1}[F'(x_0) - F'(x)]\| < q, \qquad 0 < q < 1. \tag{18}$$

这时,

$$K(x) = x - A^{-1}F(x),$$

从而 $K'(x)$ 在假定的条件下存在,且等于[3],[4]

$$K'(x) = I - A^{-1}F'(x),$$

其中 $I$ 为 $X$ 中的不变运算子. 于是依(18),

$$\alpha \equiv \sup_{\|x-x_0\| \leqslant r} \|K'(x)\| = \sup_{\|x-x_0\| \leqslant r} \|A^{-1}[A - F'(x)]\| \leqslant q.$$

---

1) 值得注意关于 Fenyö 的定理 1, 除 Вайнберг, М. М.[8] 所指出错误地用了 Lagrange 公式于运算子的情形(这是不正确的)和 Мысовских, И. П.[6]所指出的关于 Lipschitz 条件的不精确的陈述外, Fenyö 还在证明中引用了 $\|F'(x_n)\| \leqslant \frac{\eta}{r}$,这显然是错误的,因为如此则

$$1 = \|F'(x_n)F'(x_n)^{-1}\| \leqslant \|F'(x_n)\| \cdot \|F'(x_n)^{-1}\| \leqslant \frac{3}{2} B \cdot \frac{\eta}{r} = \frac{3}{2} \cdot \frac{2 - 3BCr}{11} < \frac{3}{11},$$

得出矛盾.

于是依 Taylor 公式[见[3],第四章,IV],

$$\|K(x_1)-K(x_2)\| \leqslant \alpha\|x_1-x_2\| \leqslant q\|x_1-x_2\|.$$

所以,如设[相应于[2],公式(20)]

$$\|A^{-1}F(x_0)\| \leqslant r(1-q),$$

则

$$\|x_1-x_0\| \leqslant r(1-q),$$

从而球

$$\|x-x_1\| \leqslant \frac{q}{1-q}\|x_1-x_0\|$$

含于球

$$\|x-x_0\| \leqslant \frac{1}{1-q}\|x_1-x_0\| \leqslant r$$

中,依本文开始处的一般结果,立即推出 Fenyö 的定理 2.

**3.** 注意这里的结果比牛顿过程具有较大的灵活性. 特别,可以取一适当的有逆有界线性运算子 $A$,使

$$\alpha \equiv \sup_{x\in G}\|I-A^{-1}F'(x)\| < 1,$$

即足以引用本文开始处所述的一般结果而保证逼近过程

$$x_{n+1}=x_n-A^{-1}F(x_n)$$

收敛于(1)的解. (17)只是 $A$ 的一种取法,但不是唯一可能的取法.

**4.** 我们还可以从上面的一般结果导出 Stein 的结果([7],(定理 1)). 改用我们的符号, Stein 假定了(8),并且设对于一切 $x_1, x_2\in S(x_0;r)$,

$$\|F(x_1)-F(x_2)-F'(x_1)(x_1-x_2)\| \leqslant L\|x_1-x_2\|,$$

而对于 $x\in S(x_0;r)$,

$$\|F(x)\| \leqslant N,$$

但存在常数 $\gamma$,使 $0<\gamma<r$,而

$$\|F(x_0)\| \leqslant (1-\alpha)\frac{\gamma}{B},$$

其中设(9)不仅对 $x_0$ 真,且对于一切 $x\in S(x;r)$ 真. 又设

$$\alpha \equiv BL+NC<1.$$

与在 1. 中一样,可以证明对于 $x_1, x_2\in S(x_0;r)$,如采用(15),则

$$\begin{aligned}\|K(x_1)-K(x_0)\| &\leqslant \|F'(x_1)^{-1}[F(x_1)-F(x_2)-F'(x_1)(x_1-x_2)]\|\\ &\quad+\|F'(x_1)^{-1}-F'(x_2)^{-1}\|\cdot\|F(x_2)\|\\ &\leqslant (BL+CN)\|x_1-x_2\|=\alpha\|x_1-x_2\|,\end{aligned}$$

而因

$$\begin{aligned}\frac{1}{1-\alpha}\|x_1-x_2\| &= \frac{1}{1-\alpha}\|x_0-F'(x_0)^{-1}F(x_0)-x_0\|\\ &\leqslant B\cdot\|F(x_0)\|\cdot\frac{1}{1-\alpha}<\gamma<r,\end{aligned}$$

所以球（$F$）确位于球 $S(x_0; r)$ 中. 于是一切所需条件满足，而这正是Stein所考虑的情形，这时牛顿程序的收敛于是得证.

**后记**：卢文先生指出，不引用 Fenyö，也可以直接证明(14). 事实上，依[5]汉译本158 页的公式(§20)，并引用 (8)，(9)，(10). 对于 $x \in S(x_0; r)$，

$$\|F'(x)^{-1} - F'(x_0)^{-1}\| \leqslant \frac{\|F'(x_0)^{-1}\|^2 \cdot \|F'(x_0) - F'(x)\|}{1 - \|F'(x_0)^{-1}\| \cdot \|F'(x_0) - F'(x)\|}$$

$$\leqslant \frac{B^2 C\|x - x_0\|}{1 - BC\|x - x_0\|} \leqslant \frac{B^2 C r}{1 - BCr}$$

$$\leqslant \frac{\frac{1}{3} B}{1 - \frac{1}{3}} = \frac{1}{2} B,$$

从而

$$\|F'(x)^{-1}\| \leqslant \|F'(x_0)^{-1}\| + \frac{1}{2} B \leqslant \frac{3}{2} B,$$

这正是(14).

## 参 考 文 献


[1] Collatz, L.: Das vereinfachte Nawtonsche Verfahren bei algebraischen und transzen denten Gleichungen, Zeitschr. *angew. Math. Mech.*, **34** (1954), 70—71.

[2] Fenyö, I.: Über die Lösung der in Banachschen Raume definierten nicht-linearen Gleichungen, *Acta. Math. Acad. Sci. Hung.*, **5** (1954), 85—93.

[3] Канторович, Л. В.: Функциональный анализ и прикладная математика, *Успехи Матем. Наук* **3:6** (1948), 89—185.

[4] Канторович, Л. В.: О методе Ньютона, *Труды Матем. Ин-та им. В. А. Стеклова*, том XXVIII (1949), 101—144.

[5] Люстерник, Л. А. и Соболев, В. И.: Элементы Функционального анализа, Гостехиздат, М.-Л. 1951 (汉译本，杨从仁译，泛函数分析概要，科学出版社).

[6] Мысовских, И. П.: Реферат 1479, Реферативный журнал, *Математика*, 1955.

[7] Stein, M. L.: On methods for obtaining solutions of fixed endpoint problems in the calculus of variations, *J. Res. Nat. Bur. Standards*, **50** (1953), 277—297.

[8] Вайнберг, М. М.: Реферат 348, Реферативный журнал, *Математика*, 1955.

# 张弛问题的最速下降法*

§1. **引言** 张弛方法对于解决如下的问题是一个极重要的方法：代数方程，微分方程的界值问题，特征值问题等. R. V. Southwell, L. Fox 及其他人用这方法解决了一些重要的实用问题. Temple 证明在为一般实用问题所满足的条件下，张弛方法实际地给出正确的解答. 他在他的论文[7]里考虑了两个方法：逐步张弛法与最速下降法. 前者的优点在于每步的计算简单；但有两个严重缺点：在很多情况 (Fox 称为劣调节)，收敛太慢，且在希尔柏特空间的情形收敛性没有保证. 后者的缺点在于每步的计算过于麻烦.

本文的目的在于导出一个兼有这两种方法之长处的方法. 我们的"最速下降"与 Temple 所用的意义上略有不同，我们的术语在实际计算上具有更多意义.

我们仿照 Л. В. Канторович 的处理方法[3]，在希尔柏特空间的情形考虑这一方法，并证明其收敛，找出其收敛速率来.

§2. 设 $A$ 是可分希尔柏特空间 $\mathfrak{H}$ 上一个有界正定对称线性运算子，即设存在正数 $m, M$，使 $0 < m \leqslant M < +\infty$，并且对于一切 $x \in \mathfrak{H}$,

$$m(x, x) \leqslant (Ax, x) \leqslant M(x, x).$$

我们的问题是求解线性方程

$$Ax = b, \tag{1}$$

其中 $b \in \mathfrak{H}$ 是已知元，$x$ 是未知元.

对应每一元 $x \in \mathfrak{H}$，有一名为残量的元

$$c = Ax - b \in \mathfrak{H}.$$

我们的方法是逐步使残量减弱(即使其范数逐步减少)，并且要使它减弱得充分快.

我们知道1)，方程(1)的解 $x^*$ 使汎函数

$$W(x) = (Ax, x) - (x, b) - (b, x)$$

达到其极小，而反之，凡使 $W(x)$ 达到其极小值的元必是方程(1)的解. 因而为了逐步接近方程(1)的解，我们应当逐步使 $W(x)$ 减小.

为方便起见，我们只考察实希尔柏特空间的情形：这时数积 $(x, y) = (y, x)$ 是实数. 我们取定 $\mathfrak{H}$ 中一个完全正交规格化组 $\{e_1, e_2, \cdots, e_n, \cdots\}$:

$$(e_i, e_k) = \begin{cases} 0, & \text{如果 } i \neq k, \\ 1, & \text{如果 } i = k. \end{cases}$$

设 $H$ 表示 $\mathfrak{H}$ 中由有穷多个 $e_i$ 张成的闭线性子空间(以后简称子空间)，而 $P$ 表示在这子空间上的投影运算子. 为简便起见，我们先考察特殊的子空间，由 $e_1, e_2, \cdots, e_n$ 张成.

我们从某一最初近似解(第 0 近似) $x_0$ 出发，在其上添加一个属于子空间 $H$ 中的量 $\varepsilon z$ ($\varepsilon$ 为实数)，使 $W(x)$ 减小得最多；更精确地说，我们择定一个与 $z$ 相应的 $\varepsilon$，使

---

* 原载：数学学报，5: 4(1955)，497—503. 本文合作者：卢文.

1) 见[5]，69 页，定理1.

$$W(x_0+\varepsilon z)-W(x_0)$$

尽可能小，并再择定方向 $z$（$\varepsilon$ 是 $z$ 的函数，从而也完全定出），使上式达到极小．用 $c_0$ 表示与 $x_0$相应的残量· $c_0=Ax_0-b$．由于

$$\begin{aligned}W(x_0+\varepsilon z)-W(x_0)&=(A(x_0+\varepsilon z),x_0+\varepsilon z)-2(x_0+\varepsilon z,b)\\&\quad-(A_0x_0,x_0)+2(x_0,b)\\&=\varepsilon^2(Az,z)+2\varepsilon(c_0,z)\\&=(Az,z)\left(\varepsilon+\frac{(c_0,z)}{(Az,z)}\right)^2-\frac{(c_0,z)^2}{(Az,z)},\end{aligned}\tag{2}$$

所以应当取

$$\varepsilon=-\frac{(c_0,z)}{(Az,z)},\tag{3}$$

并取 $z\in H$，使

$$\frac{(c_0,z)^2}{(Az,z)}\tag{4}$$

为极大．因为 $z$ 取自 $n$ 维空间 $H$，它实际上只依从于 $n$ 个实参数，所以可以用普通数学分析中求极大极小的方法．取(4)的依 $z$ 的微分，得极大的条件

$$\frac{(Az,z)\cdot 2(c_0,z)(c_0,dz)-(c_0,z)^2[(Az,dz)+(Adz,z)]}{(Az,z)^2}=0.$$

留意 $A$ 的对称性，并由(3)可知应取 $z$ 使 $(c_0,z)\neq 0$，经过简化，上面等式变成

$$(Az,z)(c_0,dz)-(c_0,z)(Az,dz)=0.$$

由于 $z\in H$，可以分别取 $dz$ 为沿 $e_1,\cdots,e_n$ 的方向，于是得

$$\frac{(c_0,z)}{(Az,z)}=\frac{(c_0,e_i)}{(Az,e_i)}\qquad(1\leqslant i\leqslant n).\tag{5}$$

于是其次一个近似解（第 1 近似）$x_1$ 乃是

$$x_1=x_0+\varepsilon z=x_0-\frac{(c_0,z)}{(Az,z)}z,\tag{6}$$

而由于

$$PAz=\sum_{i=1}^{n}(Az,e_i)e_i=\sum_{i=1}^{n}\frac{(Az,z)}{(c_0,z)}(c_0,e_i)e_i=\frac{(Az,z)}{(c_0,z)}Pc_0,$$

所以

$$z=\frac{(Az,z)}{(c_0,z)}A_H^{-1}Pc_0,\tag{7}$$

其中 $A_H$表示由 $A$ 在子空间 $H$ 上决定的一个运算子，$A_H\equiv PAP$.

用 $c_1$ 表示与新近似解 $x_1$ 相应的残量，那末

$$c_1=Ax_1-b=c_0-\frac{(c_0,z)}{(Az,z)}Az,\tag{8}$$

从而由(5)，(8)可知

$$(c_1,e_i)=0\qquad(1\leqslant i\leqslant n).$$

这就是说，$W(x)$ 在 $H$ 中的最速下降与把 $c_0$ 在 $H$ 中的分量（或投影）取消（或张弛）一点没