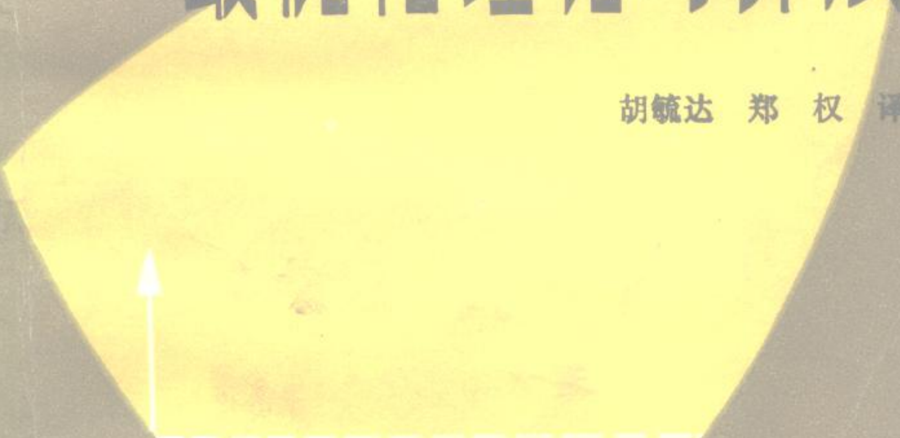
[法] J.Céa 著

最优化理论与算法

胡毓达　郑权　译



高等教育出版社

[法] J. Céa 原著

最优化理论与算法

胡毓达 郑 权 译

高等教育出版社

出版前言

本书根据法国 Dunod 出版社 1971 年出版的 J. Céa 著 Optimisation: théorie et algorithmes 一书译出．它利用泛函分析的工具介绍了最优化的基本理论和一些重要算法，还涉及了某些较深入的问题．为了方便没有学过泛函分析的读者，专门用前两章叙述这方面必要的知识．

本书可供高等学校数学专业或应用数学专业教师、研究生和高年级学生阅读，也可供从事最优化理论或算法研究的人员参考．

[法] J. Céa 原著

最优化理论与算法

胡毓达 郑 权 译

*

高等教育出版社出版

新华书店北京发行所发行

河北香河县印刷厂印装

*

开本 850×1168 1/32 印张 8.625 字数 206,000

1982 年 11 月第 1 版 1983 年 10 月第 1 次印刷

印数 00,001-- 8,500

书号 13010·0829 定价 1.30 元

译　序

最优化是近代应用数学的一个新的分支．最优化，主要是研究在给定的条件下，如何作出最好的决策去完成所与的任务，它的主要内容，就是寻求达到最好决策的方法以及建立这些方法所依据的理论．最优化是密切结合实际的需要而发展起来的，它在工程设计，计划管理，科学试验以及军事科学等方面都有着广泛的应用．

近代的最优化方法是从1947年线性规划方法的出现开始的．在这之前，研究最优化的方法是古典的微分法和变分法．那时，由于这些方法本身的复杂性以及计算工具的限制，最优化没有得到发展．在五十年代，鉴于线性规划的实际效果，再加上大规模生产的需要，使得最优化的研究蓬勃兴起．六十年代，由于与高速电子计算机结合，出现了大量最优化的算法，同时，逐渐形成了最优化的基本理论．七十年代是最优化的兴旺发展时期，这时，与最优化有关的著作、杂志和专门研究机构大量地涌现．在这一时期，最优化不论在理论或方法上又都有了新的进展．目前，国内外许多大学都已设有最优化课程和研究机构．二十多年来最优化的发展情况表明，它是一门应用很广，生气勃勃的新兴学科，并且在它的内部也出现了许多新的分支．

在我国，四十年代后期就有人从事最优化的研究工作．近二十年来，随着电子计算机的日益广泛使用，特别是在工程设计中，最优化的方法得到了越来越多的应用．最近十年，我国的数学和运筹学工作者以及广大科技工作者，无论在最优化的理论方面还是在应用方面都取得了不少可喜的成果．中国科学院有关研究所和不少大学已招收过最优化方向的研究生；一些大学的数学系或应用数学系已开始设置了最优化专门化；上百所大学的数学专业

和工程专业开设了有关最优化的各种选修课；为工程技术人员和管理人员开办的最优化学习班，也举行过许多次．当前，我国最优化的研究和应用队伍正在不断地壮大．

本书以较短的篇幅，介绍了最优化的基本理论和一些重要的算法，还涉及了一些较深入的问题．它是作者在几次教学实践的基础上写成的，所以，也是一本较为成熟的著作．对于数学基础较好的读者，特别是初步掌握了泛函分析知识的读者，通过这本书的学习，可以较快地掌握最优化理论的主要思想与方法．现在，我们把它介绍给大家，希望能对我国在最优化方面的研究和教学工作起积极的作用．

本书的第一章和第二章是以后三章的数学基础：第一章阐述与最优化理论有关的 Banach 空间和 Hilbert 空间的知识，第二章介绍按 Gateaux 意义和 Fréchet 意义的求导．第三章研究无约束最优化，介绍了一些比较成熟的算法．第四章讨论有约束最优化，主要研究了具有凸约束的凸泛函的最优化问题．最后，第五章，利用极小-极大定理和 Hahn-Banach 定理简要地讨论了对偶性．

书末转录原书所附作了分类的参考书目录．这些著作是学习最优化的很好的资料．

在翻译本书的过程中，对所发现的原书的不少错处已作了改正，并都不加译注．南京大学何旭初教授仔细地阅读了全部译稿，并提出了宝贵的意见．华东师范大学程其襄教授和复旦大学夏道行教授为个别译名提供了建设性的意见．中国科学院应用数学研究所越民义教授关心本书的翻译出版，还特别为这篇‘译序’作了具体的修改和补充．译者在此一并表示衷心的感谢！限于译者的水平，译书中难免有不妥之处，切望读者批评指正．

胡毓达　郑　权

1982年5月1日于上海

序　言

最优化是日常生活中一个很自然的概念：对一个给定的问题可以有许多可能的解答，人们要选择(一般地说)一个被认为"最优的"解答．从数学观点来看，最优化问题是，给出一个集合U和一个定义在U上并在$\boldsymbol{R}$中取值的函数J，要寻找这样的u，使得

$$\begin{cases} u \in U, \\ J(u) \leqslant J(v), \ \forall v \in U, \end{cases} \tag{*}$$

其中U表示为了达到某种目的或使某一系统得以运行的各个可能的解答的集合．在实际中，一事物是否属于U通常是用此事物是否服从一些约束来说明的，因此通常称此U为约束区域(或者称它为使约束成立的点的集合)．J表示一个用以对解进行选择的准则．问题(*)就是要选择一个解u，使得J的值达到集合U上的极小．如果把J改成$-J$，则问题转变成求极大的问题．在实际问题中，J可代表成本，工作效率，利润，持续时间，等等．

应该指出，问题(*)可能无解，在这种情况下，就是要寻找一个使$J(u)$能"充分接近"于$\operatorname{Inf}_{v \in U} J(v)$的可实现的解$u$．

集合U通常是用一些约束系统中各种参数的等式和不等式来描述的．当一个系统相当复杂时，对它进行描述就变得十分困难，即使是用数值方法也是如此．于是，就得利用一种对实际对象的描述作了简化的"模型"．然而，效能越来越好的数字电子计算机的问世，使我们有可能用越来越细致的复杂模型来工作．在一般情况下，要得到最优化问题的最优解的解析表达式是不可能的，因此，就得把问题转化为寻找这种解的一个数值近似的问题．研究数值方法的问题和数字电子计算机的性能是紧密联系的．

所有这些方法的基本思想如下：用一系列'小'问题替代'大'

问题．在这些小问题中一般仅含有一个变量．不过人们逐渐借助于把一个大问题化为一些中等程度的问题的分解方法（这有赖于在数字电子计算机上进行并行计算的可能性）．另一个研究方向是：由于有约束的问题很难解，所以应设法除去这些约束，这就导致罚函数法．应该指出，这只是改变了所给问题的性质的一种特殊情况．力学家，经济学家，…等提出另一种处理方式，就是把所给的问题与一个叫做对偶问题的新问题联系起来．

为了便于阅读本书，我们在第一章中给出了后面要用到的所有数学工具，主要是 Banach 空间和 Hilbert 空间，线性连续算子，Hahn-Banach 定理，强拓扑，弱拓扑和弱 * 拓扑以及紧性等概念．

第二章考察 Gateaux 和 Fréchet 意义下的求导问题．这些概念是重要的，它使我们一开始就能用与问题所在空间的维数无关的方式来讨论问题．

第三章是在无约束情况下求泛函极小的问题．这里，阐述了我们最感兴趣的一些方法．在把这些方法"统一处理"上也作了一些努力．

第四章考虑有约束问题．我们限于考虑凸的情形，因为一般情形的结果还是相当贫乏的．除了在 § 1.n°5 中对某些方法只给出相当简略的叙述外，算法的收敛性通常是予以证明的．在 § 2 和 § 3 中研究了罚函数法和分解方法．

最后，在第五章中，利用 Hahn-Banach 定理和极小-极大定理，简略地研究了对偶性，并给出相当数量的例子．

差不多本书中介绍的所有的方法都经过了数值试验．数值方法的比较是个很细致的问题，在本书中没有给予研究．为了对此问题作出部分的回答，我们推荐读者参考 A. R. Colville 先生的试验．

这本书取材于我为 Rennes 理学院（1966—1967，1967—1968），加州大学洛杉矶校(1968—1969)和法兰西电力公司-原子能委员会数值分析(夏季)讲习班(1969 年 7 月)准备的教材．我感谢这些机构的领导和大学生，按照他们的意见和建议，我改进了这本教材．

J. L. Lions 先生好意地把本书收进他所主持的丛书中，对此我深为感谢．

对于 Dunod 出版社的出色工作也表示感谢．

J. Céa

目　录

第一章　泛函分析初步

§1. Banach 空间

我们用 Λ 表示实数域或复数域. 在这一章中引进的所有空间都是 Λ 上的向量空间. 除非有必要,我们将不确切指明数域 Λ 是实的还是复的.

1. Banach 空间的定义. 简单性质

如果向量空间 V 到非负实数全体 $\boldsymbol{R}_+$ 中的映象 p 具有性质:

(N1)　$p(\lambda v)=|\lambda|p(v)\quad \forall v\in V,\ \forall\lambda\in\Lambda,$

(N2)　$p(u+v)\leqslant p(u)+p(v)\quad \forall u,\ v\in V,$

则称它是 V 上的半范数. 如果除了条件(N1),(N2)外,还满足条件:

(N3)　$v\in V,\ p(v)=0\Rightarrow v=0,$

则称 p 是 V 上的范数。由(N1),条件(N3)可改为

$$v\in V,\ p(v)=0\Longleftrightarrow v=0.$$

通常,用 $\|\cdot\|_V$ 或 $\|\cdot\|$ 来表示 V 上的范数. 给定了范数的向量空间叫做赋范向量空间. 赋范向量空间中两个元素 u 和 v 之间的距离用公式

$$d(u,\ v)=\|v-u\|$$

来定义. 使得

$$d(u,v)<r$$

成立的元素 $v\in V$ 的集合叫做以点 u 为中心 r 为半径的开球.

在赋范向量空间中用如下方式定义拓扑: 点 $u\in V$ 的任意一

个邻域是一个以点 u 为中心的半径非零的开球. 此外，开集是开球的和. 收敛的概念如下：如果

$$\lim_{n\to\infty}\|v_n-v\|=0,$$

则称序列 v_n 收敛于 v，这样的收敛叫做强收敛(以区别于下面引进的其它收敛概念).

大家知道，空间 V 中的 Cauchy 序列(按强拓扑意义)是满足下列条件的序列：对于任一 $\varepsilon>0$，存在 $N(\varepsilon)$，使得

$$m,n>N(\varepsilon)\Rightarrow\|v_m-v_n\|<\varepsilon.$$

如果赋范向量空间的每一个 Cauchy 序列在其中都收敛（于这个空间的元素)，则称这个空间是完备的.

定义 1.1. 完备赋范向量空间称为 Banach 空间.

例 1.1. 空间 $\boldsymbol{R}^n$ 和 $\boldsymbol{C}^n$ (实和复的 n 维空间).

设 $x=(x_1,\cdots,x_n)\in\boldsymbol{R}^n$，令

$$\|x\|=\left(\sum_{i=1}^n x_i^2\right)^{\frac{1}{2}}.$$

大家知道，按此方式定义的范数，空间 $\boldsymbol{R}^n$ 是完备的.

设 $x=(x_1,\cdots,x_n)\in\boldsymbol{C}^n$. 具有范数

$$\|x\|=\left(\sum_{i=1}^n |x_i|^2\right)^{\frac{1}{2}}$$

的空间 $\boldsymbol{C}^n$ 也是完备的.

例 1.2. 用 $\mathscr{C}([0,1])$ 表示定义在闭区间 $[0,1]$ 上，而在 $\boldsymbol{C}$ 中取值的连续函数 $f(x)$ 全体构成的空间. 在 $\mathscr{C}$ 中定义范数

$$\|f\|=\operatorname{Sup}_{x\in[0,1]}|f(x)| \tag{1.1}$$

(验证这是范数!)，我们来证明这个空间是完备的. 设 f_n 是 $\mathscr{C}$ 中的一个 Cauchy 序列，这时

$$\begin{cases}\forall\varepsilon>0,\ \exists N(\varepsilon),\text{使得}\\ n,m>N(\varepsilon)\Rightarrow\|f_n-f_m\|<\varepsilon.\end{cases}$$

从范数的定义推出

$$\begin{cases}\forall \varepsilon>0,\ \exists N(\varepsilon),\text{使得}\\ n,\ m>N(\varepsilon)\Rightarrow |f_n(x)-f_m(x)|<\varepsilon \quad \forall x\in[0,1].\end{cases} \tag{1.2}$$

所以，对于每一个 x，数列 $f_n(x)$ 都是 Cauchy 序列，因而，存在一个数，记为 $f(x)$，使得

$$\lim_{n\to\infty} f_n(x)=f(x).$$

在(1.2)式中令 m 趋于 $+\infty$，即得

$$\begin{cases}\forall \varepsilon>0,\ \exists N(\varepsilon),\text{使得}\\ n>N(\varepsilon)\Rightarrow |f_n(x)-f(x)|<\varepsilon \quad \forall x\in[0,1].\end{cases}$$

因此，f 是$[0,1]$上连续函数序列均匀收敛的极限. 由此推出 f 是连续的，这就是说 $f\in\mathscr{C}$，并且$\mathscr{C}$是完备的. 可见，具有范数(1.1)的空间$\mathscr{C}([0,1])$是 Banach 空间.

例 1.3. 设 p 满足条件：$1<p<+\infty$. 用 l^p 表示 Λ 中满足

$$\sum_{k=0}^{\infty}|a_k|^p<+\infty$$

的元素序列 $a=(a_0,\ a_1,\ \cdots,\ a_k,\ \cdots)$的空间. 我们来证明

$$\|a\|_p=\left(\sum_{k=0}^{\infty}|a_k|^p\right)^{1/p} \tag{1.3}$$

是一种范数，此外，按此范数 l^p 是完备的.

容易验证，如果 $a,\ b\in l^p$，则 $a+b\in l^p$.

关系式(N1)显然成立. 为了验证关系式(N2)，先来证明两个引理.

引理 1.1. 设 $\alpha,\beta>0$，$p>1$，$\frac{1}{p}+\frac{1}{p'}=1$，则

$$\alpha\cdot\beta\leqslant\frac{\alpha^p}{p}+\frac{\beta^{p'}}{p'}.$$

(只需令 $\gamma=\alpha\cdot\beta^{1/(1-p)}$，并证明函数$\frac{\gamma^p}{p}+\frac{1}{p'}-\gamma$ 在 $\gamma=1$ 时达到极

小即可.)

引理 1.2.(Hölder 不等式) 设 $a\in l^p$, $b\in l^{p'}$, 则

$$\sum_{k=0}^{\infty}|a_k\cdot b_k|\leqslant\|a\|_p\cdot\|b\|_{p'}. \tag{1.4}$$

证. 令

$$\alpha_k=\frac{|a_k|}{\|a\|_p},\quad \beta_k=\frac{|b_k|}{\|b\|_{p'}},$$

由引理 1.1 可得

$$\frac{|a_kb_k|}{\|a\|_p\cdot\|b\|_{p'}}\leqslant\frac{|a_k|^p}{p\|a\|_p^p}+\frac{|b_k|^{p'}}{p'\|b\|_{p'}^{p'}}.$$

对上式求和得

$$\sum_{k=0}^{\infty}\frac{|a_kb_k|}{\|a\|_p\cdot\|b\|_{p'}}\leqslant\frac{1}{p}+\frac{1}{p'}=1.$$

现在来建立(N2). 设 $a,b\in l^p$,则

$$\|a+b\|_p^p=\sum_{k=0}^{\infty}|a_k+b_k|^p\leqslant\sum_{k=0}^{\infty}|a_k+b_k|^{p-1}\cdot|a_k|+\sum_{k=0}^{\infty}|a_k+b_k|^{p-1}|b_k|.$$

令 $c_k=|a_k+b_k|^{p-1}$,因为

$$\begin{aligned}\|c\|_{p'}&=\left(\sum_{k=0}^{\infty}|a_k+b_k|^{(p-1)p'}\right)^{1/p'}\\&=\left(\sum_{k=0}^{\infty}|a_k+b_k|^p\right)^{1/p'}=\|a+b\|^{p/p'},\end{aligned}$$

故 $c\in l^{p'}$.

其次,应用引理 1.1 两次可得

$$\|a+b\|_p^p\leqslant\|c\|_{p'}\cdot\|a\|_p+\|c\|_{p'}\cdot\|b\|_p\leqslant\|a+b\|^{p/p'}(\|a\|_p+\|b\|_p),$$

从而

$$\|a+b\|_p\leqslant\|a\|_p+\|b\|_p \tag{1.5}$$

(这就是所谓 Minkowski 不等式).

现在来证明 l^p 是完备的. 设 $a^n=(a_0^n,\cdots,a_k^n,\cdots)$是 l^p 中某个 Cauchy 序列,这时

$$\begin{cases}\forall \varepsilon>0,\ \exists N(\varepsilon),\text{使得}\\ n,m>N(\varepsilon)\Rightarrow\|a^n-a^m\|<\varepsilon.\end{cases}\tag{1.6}$$

对于指定的 k，数列 a_k^n 是 Cauchy 序列，因而存在数 a_k，使得

$$\lim_{n\to\infty} a_k^n = a_k.$$

其次，(1.6)可写为如下形式：

$$\begin{cases}\forall \varepsilon>0,\ \exists N(\varepsilon),\text{使得}\\ n,\ m>N(\varepsilon)\Rightarrow\sum_{k=0}^{\infty}|a_k^n-a_k^m|^p<\varepsilon^p.\end{cases}$$

特别地，如果 q 表示某一个正整数，则

$$\begin{cases}\forall \varepsilon>0,\ \exists N(\varepsilon),\text{使得}\\ n,m>N(\varepsilon)\Rightarrow\sum_{k=0}^{q}|a_k^n-a_k^m|^p<\varepsilon^p.\end{cases}$$

令 m 趋于无穷，再令 q 趋于无穷，即得

$$n>N(\varepsilon)\Rightarrow\sum_{k=0}^{\infty}|a_k^n-a_k|^p<\varepsilon^p,$$

由此易知

$$\begin{cases}a\in l^p,\\ \lim\limits_{n\to\infty} a^n=a\ (\text{在 } l^p \text{ 中}).\end{cases}$$

可见，按照(1.3)定义的范数，空间 l^p 是 Banach 空间. ▮

例 1.3′. 空间 l^p 在 $p=1$, $p=+\infty$ 情形.

对于 $a\in l^1$，有

$$\|a\|_1=\sum_{k=0}^{\infty}|a_k|.$$

$a\in l^\infty$ 的意义是序列 a_k 有界，这时令

$$\|a\|_\infty=\operatorname*{Sup}_{k=0,\cdots,\infty}|a_k|.$$

容易验证，按照刚才引进的范数，l^1 和 l^∞ 都是 Banach 空间.

例 1.4. 空间 $L^p(\Omega)$（或 $L_p(\Omega)$, $1\leqslant p<+\infty$），其中 Ω 表示