

熵与最优化方法

ENTROPY AND OPTIMIZATION METHODS

李兴斯 姜昱汐 潘少华 © 著



大连理工大学出版社
Dalian University of Technology Press

熵与最优化方法

ENTROPY AND OPTIMIZATION METHODS

李兴斯 姜昱汐 潘少华 © 著



大连理工大学出版社
Dalian University of Technology Press

图书在版编目(CIP)数据

熵与最优化方法 / 李兴斯, 姜昱汐, 潘少华著. —
大连: 大连理工大学出版社, 2016.9

ISBN 978-7-5685-0605-2

I. ①熵… II. ①李… ②姜… ③潘… III. ①熵—最
优化算法 IV. ①O414.11

中国版本图书馆 CIP 数据核字(2016)第 246119 号

大连理工大学出版社出版

地址:大连市软件园路 80 号 邮政编码:116023

发行:0411-84708842 邮购:0411-84703636 传真:0411-84701466

E-mail: dulp@dulp.cn URL: <http://www.dulp.cn>

大连日升彩色印刷有限公司 大连理工大学出版社发行

幅面尺寸: 147mm×210mm	印张: 9.5	字数: 193 千字
2016 年 9 月第 1 版	2016 年 9 月第 1 次印刷	

责任编辑: 王伟

责任校对: 苇杭

封面设计: 王薇薇

ISBN 978-7-5685-0605-2

定 价: 38.00 元

本书由大连市人民政府资助出版

**The published book is sponsored
by the Dalian Municipal Government**

前 言

本书主要总结和介绍了我们科研团队多年来关于最优化方法的研究成果,内容涉及线性规划、非线性规划、离散优化和不可微优化中的一些重要问题。有别于一般数学规划类的书籍和文献,我们解决这些问题的方法借助了一个特殊的工具,即书名中的“熵”。为方便读者阅读和理解,将全书划分为上、中、下三篇。

上篇(第1,2章)主要介绍熵的基本概念和两个相关的优化原理。我们沿着其发展历程,简要介绍了熵在热力学、统计物理学和信息论中的含义,而其作为无序性或不确定性程度的普适定义,使其具备了跨越自然科学和社会科学各学科领域的特质。最大熵原理与最小叉熵原理包含概率测度,使其成为熵函数得以广泛应用的有效载体。

中篇(第3~6章)讨论的熵与最优化方法是本书的核心部分。第3章针对不可微的有限极大极小问题,通过在其拉格朗日函数中引进熵函数的正则化项,推导出可以一致逼近极大值函数的光滑函数,即所谓的凝聚函数,为解决一大类与极大值函数相关的不可微问题扫除了障碍。第4章利用凸分析知识,从正则化的角度为熵函数方法寻找理论依据,证明了通过熵正则化法导出的光滑

函数即为传统的指数(乘子)罚函数。尽管 Ben-Tal 等人使用回收函数也导出了同样的光滑函数,但他们的方法需要较深的凸分析知识。第 5 章将熵正则化方法的思想延伸到约束优化问题,通过在经典拉格朗日函数中引进正则化项,不仅可以导出传统的指数(乘子)罚函数,而且可以通过选择不同的正则化函数,导出其他惩罚(障碍)函数。这种拉格朗日正则化方法在两种求解约束优化的无约束方法(拉格朗日函数法与惩罚函数法)之间搭建了一个桥梁,不仅给出了构造惩罚函数的一个统一框架,而且赋予惩罚函数一种新的解释。Auslender 使用回收函数方法也给出了构造惩罚函数的一般方法,这与前面提到的 Ben-Tal 等人的工作一样,与我们的正则化方法可谓殊途同归。第 6 章通过用 D -函数和 φ -散度代替叉熵函数,给出求解凸规划的 Bregman 邻近点算法和熵型邻近点算法。

下篇(第 7~9 章)回归到熵函数法在传统数学规划问题中的应用。第 7 章针对线性规划内点法使用的摄动方程进行了分析研究,提出了三种改造中心化方程的方法,其目的是改善现有内点法在中心化效率方面的不足,以加强其缩减对偶间隙的力度。第 8 章讨论非线性规划,首先针对不可微的精确罚函数,利用凝聚函数建立了相应的光滑函数形式,即所谓的“准”精确罚函数。因为一个优化问题的所有约束可以用一个极大值约束等价代替,故可将一个多约束优化问题转化为一个单约束优化问题,进而利用凝聚函数光滑化极大值函数形式的单约束。第 9 章对 0-1 规划等离散优化问题,建立了相应的连续化算法。

感谢国家自然科学基金(71301015)与大连市政府学术专著出

版基金对本项目的资助,同时感谢大连理工大学出版社在专著出版过程中的大力帮助。

本专著所进行的研究工作,得到已故中国科学院资深院士钱令希先生的长期关心和支持,值此恩师百年华诞之际,仅以专著的出版告慰恩师在天之灵。同时,感谢钟万勰院士和程耿东院士对该项研究工作的支持。

鉴于作者水平所限,书中定有错误或不妥之处,敬请读者指正。

李兴斯

2016年6月

目 录

上篇 熵与熵优化原理

第 1 章 熵的基本概念 / 3

- 1.1 引 言 / 3
- 1.2 热力学中的熵 / 4
- 1.3 统计物理学中的熵 / 6
- 1.4 信息论中的熵 / 9
- 1.5 小 结 / 11

第 2 章 熵优化原理及其对偶 / 12

- 2.1 引 言 / 12
- 2.2 最大熵原理 / 13
- 2.3 最小叉熵原理 / 17
- 2.4 熵优化问题的对偶规划 / 20
- 2.5 熵优化问题与几何规划的关系 / 24
- 2.6 小 结 / 27

中篇 熵与最优化方法

第3章 有限极大极小问题的熵函数方法 / 31

3.1 引言 / 31

3.2 熵函数方法 / 33

3.2.1 极大熵方法 / 33

3.2.2 最小叉熵函数方法 / 38

3.3 光滑函数 $F_p(x)$ 与 $F_p(x, \mu)$ 的性质 / 40

3.4 光滑化算法及其收敛性分析 / 45

3.5 光滑化算法的实现及数值计算结果 / 50

3.5.1 $F_p(x)$ 与 $F_p(x, \mu)$ 的函数值及其梯度的计算 / 50

3.5.2 算法实现及计算结果 / 52

3.6 小结 / 56

第4章 熵函数方法与传统优化方法的关系 / 57

4.1 引言 / 58

4.2 熵函数方法的正则化理论 / 60

4.3 熵函数方法与指数(乘子)罚函数方法之间的关系 / 63

4.3.1 求解有限极大极小问题的指数(乘子)罚函数方法 / 64

4.3.2 熵函数方法与指数(乘子)罚函数方法

之间的对偶关系 / 69

4.4 熵函数方法与 Ben Tal 等人的光滑化法的关系 / 71

4.5 小结 / 73

第 5 章	约束优化问题的拉格朗日正则化方法	/ 76
5.1	引 言	/ 77
5.2	约束优化问题的熵正则化方法	/ 79
5.3	拉格朗日正则化方法	/ 83
5.4	拉格朗日正则化方法的收敛性分析	/ 85
5.5	构造罚函数的统一框架与实例	/ 88
5.5.1	构造罚函数的统一框架	/ 88
5.5.2	构造罚函数的实例	/ 91
5.6	小 结	/ 96
第 6 章	解凸规划的邻近点算法	/ 98
6.1	引 言	/ 98
6.2	凸规划 Bregman 邻近点算法	/ 101
6.2.1	Bregman 函数、 D -函数和 D_h 投影	/ 102
6.2.2	D -函数邻近极小化算法	/ 107
6.2.3	乘子法	/ 110
6.2.4	邻近乘子法	/ 117
6.3	凸规划的熵型邻近点算法	/ 121
6.3.1	φ -散度度量	/ 122
6.3.2	熵型邻近点算法	/ 125
6.3.3	熵型乘子法	/ 138
6.3.4	熵型邻近点算法的具体应用与实例	/ 142
6.4	小 结	/ 148

下篇 熵函数方法在数学规划中的应用

第7章 解线性规划的原-对偶路径跟踪算法 / 153

7.1 引言 / 154

7.2 基于代数等价变换的原-对偶路径跟踪内点算法 / 158

7.2.1 邻近性度量的概念 / 158

7.2.2 代数等价变换与相应的邻近性度量和 搜索方向 / 160

7.2.3 基于代数等价变换的不可行原-对偶路径 跟踪内点算法 / 165

7.2.4 算法的收敛性与多项式复杂性界限分析 / 167

7.3 自调节原-对偶路径跟踪内点算法 / 177

7.3.1 一个新的邻近性度量函数 / 178

7.3.2 具有自调节功能的原-对偶路径 跟踪内点算法 / 179

7.3.3 算法实现及计算结果 / 183

7.4 一个非内点原-对偶路径跟踪算法 / 190

7.4.1 非内点原-对偶路径跟踪算法的基本思想 / 191

7.4.2 非内点原-对偶路径跟踪算法的实现 / 195

7.4.3 算法的收敛性分析 / 198

7.4.4 算法实现及计算结果 / 208

7.5 小结 / 210

第 8 章 熵函数方法在非线形规划中的应用 / 212

8.1 引 言 / 213

8.2 可微“准”精确罚函数方法 / 214

8.2.1 “准”精确罚函数方法的主要思想 / 214

8.2.2 “准”精确罚函数方法的基本理论 / 216

8.2.3 基本算法与数值算例 / 225

8.3 凝聚函数方法 / 227

8.3.1 凝聚函数方法的基本思想 / 227

8.3.2 基本算法与数值算例 / 229

8.4 小 结 / 231

第 9 章 离散优化的连续化方法 / 233

9.1 求解线性 0-1 规划的一种连续化方法 / 233

9.1.1 基于拉格朗日松弛的连续化方法 / 234

9.1.2 算法实现 / 235

9.2 二进制二次规划的连续化方法 / 238

9.2.1 引 言 / 238

9.2.2 基于 NCP 函数的连续优化模型 / 239

9.2.3 二进制二次规划问题的全局连续化算法
与收敛分析 / 242

9.2.4 算法实现及计算结果 / 246

9.3 最大函数之和极小化的一个连续化方法 / 251

9.3.1 引 言 / 252

9.3.2 组合优化问题的连续化模型与光滑近似 / 253

9.3.3 算法实现及计算结果 / 254

9.4 小 结 / 257

参考文献 / 259

附 录 / 274

上 篇

熵与熵优化原理

主要介绍熵的基本概念和两个相关的优化原理。我们沿着熵的发展历程,简要介绍了熵在热力学、统计物理和信息论中的含义。而其作为无序性或不确定性程度的普适定义,使其具备了跨越自然科学和社会科学各学科领域的特质。最大熵原理与最小叉熵原理包含着概率测度,这使其成为熵函数得以广泛应用的有效载体。

第 1 章 熵的基本概念

本章主要向读者介绍热力学、统计物理和信息论等学科中熵的含义。

1.1 引 言

熵的概念最初源自于热力学,后来进一步发展到统计物理与信息论等学科。在热力学、统计物理和信息论等学科领域里,熵的概念占据着一个中心位置。在热力学里,它作为热力学第二定律的基础,指示着一个热力学过程得以发生的方向,并同时作为一个孤立系统达到热力学平衡态的判据。在统计物理中,它作为沟通热力学宏观态与微观态之间的桥梁,建立起热力学宏观物理量与微观物理量之间的定量关系。在信息论里,它作为现代信息论的基础,建立了关于信息的某种度量。下面就把熵在这几个学科里的不同含义分别予以介绍。但因为本书的主旨是介绍以熵为基础的最优化算法,同时熵的概念又横跨诸多学科,所以不可能对其在不同学科里的含义,通过短短的几页书做完整而精确的介绍,有的内容还按个人理解做了一些必要的简化。希望对有关概念详加了解的读者,可参看有关教材。

1.2 热力学中的熵

热力学中关于熵的概念,最早是由德国科学家 R. Clausius^[1]在 1865 年前后引进的。它与内能一起作为热力学的两个重要状态函数,分别成为热力学第二和第一定律的基础。而且,从它在热力学的中所起的作用来看,其重要性甚至还要超过内能。关于能量与熵这两个物理量在热力学中所起的不同作用,下面 Emden 的一段话做了非常精辟的说明^[2]。他说:“在自然过程的巨大工厂里,熵的原理占据着经理的位置,因为它掌握着整个企业的运作方式;而能量原理仅起一个平衡信贷的簿记作用。”为了使读者更精确地理解这段话,特将英文原话放在脚注里^①。下面通过热力学的两个最基本定律,来具体说明能量与熵在热力学中的这两种不同作用。

热力学第一定律被叙述为:存在一个被称之为内能的状态函数 $U(a_k, T)$,使得

$$dU = dW + dQ = \sum_k A_k da_k + dQ \quad (1-1-1)$$

其中 dU 是系统内能 U 的增量, dW 和 dQ 分别是外界对系统所做功的增量与外界供给系统的热量的增量;而 A_k 和 a_k 分别为广义力与广义位移。对于稀薄气体,因有 $dW = -p dV$,故广义力与广义位移分别对应于静水压力 ($-p$) 与体积 V 。根据这个定律,内能 U 的全微分等于对系统所做的元功与所供给的元热之和。因此,热力学第一定律是一个包括内能在内的更为广泛适用的能量守恒定律,而方

① Emden wrote: "In the huge manufactory of natural processes, the principle of entropy occupies the position of manager, for it dictates the manner and method of the whole business, whilst the principle of energy merely does the book-keeping, balancing credits and debits."