

... 大电网最优潮流计算

刘明波 谢 敏 赵维兴 / 著



科学出版社
www.sciencep.com

大电网最优潮流计算

刘明波 谢 敏 赵维兴 著

科 学 出 版 社

北 京

前 言

日益扩大的电网规模、复杂的运行方式和调控难度为电力系统运行和控制带来巨大挑战。传统的经济调度及以经济性作为主要目标的无功优化和电压调控等手段已难以适应当前大电网在经济性的基础上同时考虑安全性的要求,迫切需要发展新的电网优化运行、调度以及调控分析方法。起源于 20 世纪 60 年代的最优潮流作为电力系统最为基本,且最为重要的分析计算工具之一,已在电网经济调度、无功优化、电压调控等领域逐步获得推广应用。但目前应用于大电网的最优潮流计算不可避免地存在计算维数过高,计算量过大且求解效率偏低的问题,因此,在实际应用中,通常会对一些约束条件以及电网运行的实际情况进行相应简化后再开展最优潮流计算,以加快运算速度,提高优化求解的效率。如何在最优潮流计算中考虑更为复杂的实际运行情况,如考虑电网安全稳定运行的要求、考虑动态无功约束等,已成为最优潮流计算在大电网全局最优调控领域深入推广和应用的难点。

自 1994 年起,我和我的研究生们开始涉足电力系统最优潮流计算领域,致力于将最优潮流计算推广应用于大规模电网的研究。近年来,我们发表的一些学术论文引起了同行学者以及电力行业专业技术人员的关注,再加上多年在指导研究生和讲授研究生专业课的过程中,深切体会到电力系统飞速发展对电力专业人才培养提出了新需求,我深感有必要将 15 年来在大电网最优潮流计算领域的研究成果汇总出版,希望能够为初次踏入电力系统优化领域的研究人员起到抛砖引玉的作用,同时也能够为广大同行和相关专业技术人员提供学习参考和探讨平台。

全书共分十七章,从内容划分上可归为三部分。前七章以线性规划和非线性规划理论为起点,将内点法及其扩展算法应用到静态无功优化、动态无功优化及其并行计算中。第八至十二章介绍了扩展的最优潮流问题,包括暂态稳定约束最优潮流和静态电压稳定约束最优潮流计算。第十三至十七章介绍了我们在大电网最优潮流分解协调计算领域的研究成果。本书对所提出的每种算法从模型建立、算法实现等方面进行了详细推导;在算例分析中,不仅采用了国际通用的标准试验系统作为算例,且采用了小至 538 节点,大至 2212 节点等省级和区域电网的实际运行数据作为算例。

在内容的具体安排上,第一章介绍了非线性规划和线性规划的概念及算法,通过引出一阶最优性条件(KKT 条件)对非线性原-对偶内点法以及线性规划单纯形法和线性规划内点法的实现原理进行介绍;第二章介绍了无功优化问题模型和灵

敏度计算方法,进而介绍无功优化问题及其逐次线性规划解法;第三章为无功优化问题的非线性混合整数规划方法,着重介绍了考虑离散变量的非线性原对偶内点法;第四章为动态无功优化问题模型及其求解方法;第五章介绍了动态无功优化问题的几种解耦算法;第六章介绍了动态无功优化并行算法及其在并行计算环境下的实现;第七章为无功优化应用背景的实践,介绍了分布式电压无功全局实时优化控制系统及在实际电网中的应用;第八章为基于故障模式法的暂态能量裕度约束最优潮流计算;第九章为基于BCU法的暂态能量裕度约束最优潮流计算;第十章介绍了基于轨迹灵敏度法的暂态稳定约束发电再调度;第十一章为基于轨迹灵敏度法的暂态稳定约束最优潮流计算;第十二章介绍了静态电压稳定裕度约束无功优化计算;第十三章开始进入了大电网最优潮流的计算,并对几种典型的分解协调算法进行阐述;第十四章应用基于近似牛顿方向的多区域无功优化分解算法进行大电网的最优潮流计算;第十五章和第十六章则分别应用基于对角加边模型的多区域无功优化分解算法和基于诺顿等值的多区域无功优化分解算法进行大电网最优潮流计算;第十七章对前述几种分解算法的收敛特性和计算速度进行了比较,并归纳出影响最优潮流分解协调算法收敛性和计算速度的因素。

本书成果得到国家自然科学基金(编号:50277013、50777021和50907023)和广东省自然科学基金(编号:011648)的资助,凝聚了我们70余篇学术论文成果,也是对我们这15年来从事大电网最优潮流计算领域研究历程的一个小结。在此感谢研究生陈学军、王晓村、程莹、李健、朱春明、夏岩、黄伟、杨勇、付钢、阳曾、赖永生、李妍红、蒋健、缪楠林、曲绍杰、李贻凯、王奇、马冠雄为本书研究成果所作的贡献。目前,他们均已走上了工作岗位,大部分从事电力系统生产、设计、调度运行与管理方面的工作,本书留下了他们在华南理工大学电力系统优化与控制科研团队的成长足迹。在研究过程中,还得到华南理工大学电力学院吴捷教授,广东电网电力调度通信中心原总工程师马志强、现任总工程师温柏坚,广东电网公司生产技术部程启诚高级工程师,南方电网调度通信中心曾勇刚副总工程师,广州供电局电力调度中心钱康龄高级工程师和李芳红高级工程师的指导和帮助,在此一并表示感谢。此外,还要感谢我的家人对我的鼓励和支持。

本书不妥之处在所难免,欢迎读者批评指正。

作 者

2010年2月15日

目 录

前言

第一章 非线性规划和线性规划的求解方法	1
1.1 基础知识	1
1.1.1 最优化问题的数学描述	1
1.1.2 相关数学基础	2
1.2 非线性规划	5
1.2.1 一阶最优性条件	6
1.2.2 二阶最优性条件	7
1.2.3 非线性原对偶内点方法	8
1.3 线性规划	12
1.3.1 单纯形法	12
1.3.2 内点法	15
1.4 小结	17
参考文献	17
第二章 连续无功优化计算	18
2.1 线性规划建模	18
2.2 灵敏度系数计算	19
2.3 灵敏度系数计算中应注意的问题	20
2.4 原对偶内点法	23
2.4.1 基本原理	23
2.4.2 线性方程组的求解	26
2.4.3 迭代步长的确定及壁垒参数的修正	28
2.4.4 初始点的选择	28
2.5 计算步骤	28
2.6 算例分析	29
2.6.1 Ward & Hale 6 节点系统	30
2.6.2 IEEE 118 节点系统	32
2.6.3 某 538 节点系统	34

2.6.4 计算时间比较	35
2.7 小结	36
参考文献	37
第三章 离散无功优化计算	38
3.1 非线性混合整数规划建模	38
3.2 内嵌离散惩罚的非线性原对偶内点法	38
3.3 离散变量的处理	42
3.4 应注意的问题	44
3.4.1 迭代步长的确定和壁垒参数的修正	44
3.4.2 初始点的选择	45
3.5 计算步骤	45
3.6 修正方程的求解	46
3.7 算例分析	50
3.7.1 Ward & Hale 6 节点系统	50
3.7.2 某 538 节点系统	52
3.7.3 不同数据结构的比较	55
3.7.4 计算时间比较	56
3.8 混合整数规划问题的连续化方法	57
3.8.1 离散变量的连续化处理	57
3.8.2 二进制编码的逐位优化	58
3.8.3 算例分析	59
3.9 小结	61
参考文献	62
第四章 动态无功优化计算	63
4.1 数学模型	63
4.2 优化算法	64
4.2.1 基本原理	64
4.2.2 迭代步长的确定	69
4.2.3 罚函数的引入	70
4.2.4 收敛精度的给定	70
4.2.5 计算步骤	71
4.2.6 修正方程的求解	72

4.3 结果分析	75
4.3.1 变压器变比	76
4.3.2 电容器组无功出力	79
4.3.3 部分连续控制变量	82
4.3.4 最大潮流偏差和补偿间隙	89
4.4 动态和静态无功优化算法比较	93
4.5 与其他三种算法的比较	96
4.5.1 GAMS	96
4.5.2 GA	97
4.5.3 BARON	99
4.5.4 DICOPT	99
4.6 小结	100
参考文献	101
第五章 动态无功优化解耦算法	103
5.1 快速解耦算法一	103
5.2 快速解耦算法二	105
5.2.1 基本思想	105
5.2.2 修正方程的快速求解	106
5.3 算例分析	106
5.3.1 鹿鸣电网 14 节点系统	106
5.3.2 修改后的 IEEE 118 节点系统	110
5.4 小结	113
参考文献	114
第六章 动态无功优化并行计算	115
6.1 MPI 并行实现技术	115
6.2 并行算法及其实现	121
6.2.1 并行求解思路	121
6.2.2 MPI 并行环境下的算法实现	121
6.2.3 MPICH 的配置	122
6.2.4 并行算法实现中的几个问题	123
6.3 算例分析	124
6.4 小结	127

参考文献	128
第七章 地区电网电压无功控制	129
7.1 电压控制的基本方式	129
7.1.1 分散控制	129
7.1.2 集中控制	130
7.1.3 关联分散控制	130
7.2 分布式电压无功控制	131
7.3 变电站电压无功控制范围的整定计算	133
7.3.1 调节范围定义	133
7.3.2 整定计算原理	134
7.3.3 电压控制范围给定	135
7.3.4 无功控制范围给定	136
7.3.5 算例分析	136
7.4 小结	141
参考文献	142
第八章 基于故障模式法的暂态能量裕度约束最优潮流计算	143
8.1 常规最优潮流模型	144
8.2 TSCOPF 模型	144
8.2.1 暂态稳定计算模型	144
8.2.2 多故障 TSCOPF 模型	147
8.3 暂态能量函数和临界能量表达式	150
8.3.1 同步坐标	150
8.3.2 惯量中心坐标	151
8.4 暂态稳定裕度计算	153
8.4.1 故障切除时刻的能量	153
8.4.2 临界能量	154
8.5 灵敏度分析	160
8.6 暂态稳定裕度灵敏度的解析方法	161
8.7 计算步骤	163
8.8 算例分析	165
8.8.1 WSCC 3 机 9 节点系统	165
8.8.2 New England 10 机 39 节点系统	170

8.9 小结	175
参考文献	175
第九章 基于 BCU 法的暂态能量裕度约束最优潮流计算	177
9.1 暂态稳定裕度灵敏度分析	177
9.1.1 暂态能量裕度计算	178
9.1.2 暂态能量裕度灵敏度计算	181
9.2 BCU 法与 MOD 法对比	183
9.3 多故障 TSCOPF 计算	185
9.4 算例分析	186
9.4.1 单故障 TSCOPF 扫描结果	186
9.4.2 考虑暂态稳定约束前后 OPF 结果对比	190
9.4.3 单故障 TSCOPF 结果比较	192
9.4.4 多故障 TSCOPF 结果比较	194
9.4.5 故障分组结果	194
9.5 小结	196
参考文献	196
第十章 基于轨迹灵敏度法的暂态稳定约束发电再调度	198
10.1 电力系统机电暂态模型	198
10.1.1 发电机模型	199
10.1.2 励磁系统模型	201
10.1.3 机网接口及网络方程	204
10.2 基于改进欧拉法的暂态稳定计算	206
10.3 轨迹灵敏度分析	209
10.3.1 经典模型下的轨迹灵敏度分析	210
10.3.2 复杂模型下的轨迹灵敏度分析	213
10.4 发电机临界程度排序和原始有功转移功率计算	214
10.5 单一故障 TSCOPF 模型及最优转移功率的求解	216
10.5.1 单一故障 TSCOPF 模型	216
10.5.2 原始有功转移功率的求解	218
10.5.3 搜索最优转移功率的迭代算法	219
10.6 多故障 TSCOPF 模型及最优转移功率的求解	222
10.7 算例分析	224

10.7.1 发电机采用经典二阶模型	224
10.7.2 发电机采用四阶模型	228
10.8 小结	238
参考文献	238
第十一章 基于轨迹灵敏度法的暂态稳定约束最优潮流	240
11.1 轨迹灵敏度分析	240
11.1.1 初值计算	240
11.1.2 时域计算	243
11.2 基于轨迹灵敏度法的 TSCOPF	246
11.2.1 TSCOPF 二次规划模型及求解	247
11.2.2 多故障 TSCOPF 二次规划模型及求解	250
11.3 算例分析	253
11.3.1 单故障 TSCOPF 算例	253
11.3.2 多故障 TSCOPF 算例	264
11.3.3 与其他方法的比较	268
11.4 小结	269
参考文献	269
第十二章 静态电压稳定裕度约束无功优化计算	271
12.1 PV 曲线和电压崩溃点类型	271
12.2 用连续潮流法计算静态电压稳定极限	274
12.2.1 基本原理	274
12.2.2 修正方程式	275
12.2.3 修正方程式的预解	277
12.2.4 扩展状态变量修正值的计算	278
12.2.5 连续参数的选择	279
12.3 静态电压稳定裕度对变量的灵敏度计算	280
12.3.1 鞍结型分岔情形下的计算	280
12.3.2 极限诱导型分岔情形下的计算	281
12.4 考虑电压稳定裕度约束的无功优化计算	282
12.4.1 计算原理	282
12.4.2 算例与结果分析	284
12.5 基于 FVSI 指标的无功优化计算	291

12.5.1 快速电压稳定指标 FVSI	291
12.5.2 计算原理	293
12.5.3 算例与结果分析	293
12.6 小结	301
参考文献	301
第十三章 几种典型的分解协调算法	304
13.1 基于 PQ 分解技术的分解算法	305
13.2 基于 Benders 分解技术的分解算法	305
13.3 基于拉格朗日松弛技术的分解算法	306
13.4 基于辅助问题原理的分解算法	308
13.5 基于智能型优化的并行算法	308
13.6 基于协同进化法的分解算法	309
13.7 小结	310
参考文献	310
第十四章 基于近似牛顿方向的多区域无功优化分解算法	314
14.1 多区域系统无功优化模型	314
14.1.1 电力系统离散无功优化模型	314
14.1.2 区域分解及边界节点定义	315
14.1.3 多区域系统无功优化模型	315
14.1.4 最优化模型分解	315
14.2 引入离散处理机制的非线性原对偶内点法	316
14.3 近似牛顿方向和纯牛顿方向的定义	319
14.4 解耦的充分条件	319
14.4.1 解耦理论判据	319
14.4.2 解耦实用判据	320
14.5 不满足解耦条件时的计算方法	320
14.5.1 GMRES 算法	320
14.5.2 预处理技术	321
14.6 计算步骤	321
14.7 应注意的几个问题	322
14.7.1 GMRES(m)算法中 m 取值	322
14.7.2 罚函数的引入机制	322

14.7.3 收敛精度的确定	323
14.8 算例分析	323
14.8.1 1062 节点系统	324
14.8.2 538 节点系统	327
14.8.3 结果分析	330
14.9 小结	332
参考文献	333
第十五章 基于对角加边模型的多区域无功优化分解算法	334
15.1 区域分解	335
15.2 多区域系统离散无功优化模型	335
15.3 多区域分解算法	336
15.3.1 对角加边结构修正矩阵的形成	336
15.3.2 几种分解方案	340
15.4 算例系统	344
15.4.1 IEEE 118 节点系统	344
15.4.2 538 节点系统	345
15.4.3 1133 节点系统	345
15.5 计算结果分析	347
15.5.1 计算结果	347
15.5.2 分析与讨论	350
15.6 小结	352
参考文献	353
第十六章 基于诺顿等值的多区域无功优化分解算法	354
16.1 外部网络的静态等值	354
16.1.1 网络的划分	354
16.1.2 外部网络的等值方法	355
16.2 诺顿等值及分解算法的形成	356
16.2.1 系统的分解及诺顿等值模型	356
16.2.2 分解算法中的几个关键问题	357
16.3 计算误差分析	361
16.3.1 无功优化最优解的几种状态	361
16.3.2 误差分析	362

16.4 计算步骤	362
16.5 算例分析	363
16.5.1 236 节点系统	363
16.5.2 2212 节点系统	363
16.5.3 计算结果分析	365
16.6 小结	368
参考文献	369
第十七章 几种无功优化分解算法比较	371
17.1 算例系统	371
17.1.1 538 节点系统	371
17.1.2 708 节点系统	371
17.2 计算结果	373
17.3 各种分解算法的比较分析	374
17.4 影响计算效益的因素分析	374
17.4.1 子区域数目对计算速度的影响	375
17.4.2 最大子区域规模对计算速度的影响	376
17.5 小结	377
参考文献	377
附录	378
附录 I Ward & Hale 6 节点标准试验系统数据	378
附录 II IEEE 14 节点标准试验系统数据	379
附录 III IEEE 30 节点标准试验系统数据	381
附录 IV IEEE 118 节点标准试验系统数据	384
附录 V 某 538 节点实际系统概况	393
附录 VI 广州鹿鸣电网 14 节点系统数据	394
附录 VII WSCC 3 机 9 节点标准试验系统数据	397
附录 VIII New England 10 机 39 节点标准试验系统数据	400
附录 IX UK 20 机 100 节点试验系统接线图	405

第一章 非线性规划和线性规划的求解方法

最优化(optimization)指对于给定的实际问题,如何从众多的方案中选出最优方案。最优化问题可以追溯到古老的极值问题,但成为一门独立的学科是在 20 世纪 40 年代末,即在 1947 年,Dantzig 提出求解一般线性规划问题的单纯形法之后。作为最优化理论与方法重要分支的非线性规划与线性规划是揭示各类最优化问题的常用数学工具^[1~4]。本章将对非线性规划和线性规划问题的模型和求解方法进行阐述,为后续章节提供数学基础知识。对于上述两类最优化问题所涉及的相关定理,在论述过程中将不做详细数学论证,而以结论性和求解方法的可操作性作为本章重点。

1.1 基础知识

1.1.1 最优化问题的数学描述

最优化问题用数学方法进行描述时,包含了目标函数、决策变量和约束条件这三个要素。其中,目标函数度量解的优劣程度,如成本、收益、利润等;而决策变量为可以调节、影响优化过程的变量;约束条件指决策变量及其相互关系的约束。最优化问题的一般形式为:

$$\begin{cases} \min f(\mathbf{x}) \\ \text{s. t. } g_i(\mathbf{x}) = 0, \quad i = 1, 2, \dots, m \\ \quad \quad h_i(\mathbf{x}) \geq 0, \quad i = 1, 2, \dots, r \end{cases} \quad (1-1)$$

式中, $\mathbf{x} \in \mathbf{R}^n$ 为决策变量; $f(\mathbf{x})$ 为目标函数; $g_i(\mathbf{x}) = 0$ 为等式约束; $h_i(\mathbf{x}) \geq 0$ 为不等式约束。

例 1-1 最优化数学模型为:

$$\begin{cases} \min (x_1 - 2)^2 + (x_2 - 1)^2 \\ \text{s. t. } x_1^2 - x_2 \leq 0 \\ \quad \quad x_1 + x_2 \leq 2 \end{cases} \quad (1-2)$$

可见,目标函数为 $f(\mathbf{x}) = (x_1 - 2)^2 + (x_2 - 1)^2$, 决策变量为 $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, 不等式

约束函数写成向量形式为 $\mathbf{h}(\mathbf{x}) = \begin{bmatrix} h_1(\mathbf{x}) \\ h_2(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} -x_1^2 + x_2 \\ -x_1 - x_2 + 2 \end{bmatrix}$, 式(1-2)中的等式约

束函数 $g_i(\mathbf{x})$ 为空, 即不存在等式约束。

式(1-2)中的目标函数和部分约束函数是非线性函数, 故它所描述的是一个非线性规划(nonlinear programming, NLP) 问题。

例 1-2 生产计划问题

某化工厂在计划期内拟安排生产 I 和 II 两种产品, 已知其市场需求量和单位利润及原材料的消耗量如表 1-1 所示。应如何安排生产计划才能使该工厂的获利最多?

表 1-1 生产计划综合信息

	原材料消耗/t			产品利润/(千元/t)	产品需求量/t
	A	B	C		
产品 I	10	6	8	5	不大于 1500
产品 II	5	20	15	7	
原材料最大供应量/t	600	500	800		

设 x_1 和 x_2 分别表示产品 I 和 II 的产量, 则该问题的数学模型可写为:

$$\left\{ \begin{array}{l} \max 5x_1 + 7x_2 \\ \text{s. t. } 10x_1 + 5x_2 \leq 600 \\ 6x_1 + 20x_2 \leq 500 \\ 8x_1 + 15x_2 \leq 800 \\ x_1 \leq 1500 \\ x_1, x_2 \geq 0 \end{array} \right. \quad (1-3)$$

与例 1-1 所描述的非线性问题不同, 例 1-2 所描述的是一个线性规划(linear programming, LP) 问题, 即目标函数和约束函数均是线性函数。

1.1.2 相关数学基础

① 假设 $f(\mathbf{x})$ 为标量函数, $\mathbf{x} \in \mathbf{R}^n$, 则其梯度 $\nabla f(\mathbf{x})$ 定义为:

$$\nabla f(\mathbf{x}) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix} \quad (1-4)$$

② 假设 $f(\mathbf{x})$ 为标量函数, $\mathbf{x} \in \mathbf{R}^n$, 则其海森矩阵 $\nabla^2 f(\mathbf{x})$ 定义为:

$$\nabla^2 f(\mathbf{x}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix} \quad (1-5)$$

可见,海森矩阵是对称矩阵。

③ 假设 $\mathbf{g}(\mathbf{x}) = \begin{bmatrix} g_1(\mathbf{x}) \\ g_2(\mathbf{x}) \\ \vdots \\ g_m(\mathbf{x}) \end{bmatrix}$ 为函数向量,则其雅可比矩阵 $\nabla \mathbf{g}(\mathbf{x})$ 定义为:

$$\nabla \mathbf{g}(\mathbf{x}) = \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} & \cdots & \frac{\partial g_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_m}{\partial x_1} & \frac{\partial g_m}{\partial x_2} & \cdots & \frac{\partial g_m}{\partial x_n} \end{bmatrix} \quad (1-6)$$

④ 如果矩阵 $\mathbf{A} (\mathbf{A} \in \mathbf{R}^{n \times n})$ 满足 $(\mathbf{A}\mathbf{x}, \mathbf{x}) > 0$, 则该矩阵为正定矩阵; 如果矩阵 $\mathbf{A} (\mathbf{A} \in \mathbf{R}^{n \times n})$ 满足 $(\mathbf{A}\mathbf{x}, \mathbf{x}) \geq 0$, 则该矩阵为正半定矩阵。

⑤ 矩阵 $\mathbf{A} (\mathbf{A} \in \mathbf{R}^{n \times n})$ 的特征值 λ_i 定义为:

$$\det(\mathbf{A} - \lambda_i \mathbf{I}) = 0, \quad \mathbf{A} \in \mathbf{R}^{n \times n} \quad (1-7)$$

如果 $\mathbf{A}$ 是对称的, 则所有特征值 λ_i 为实数。当 $\lambda_i > 0, i=1, 2, \dots, n$ 时, $\mathbf{A}$ 是对称正定的; 当 $\lambda_i < 0, i=1, 2, \dots, n$ 时, $\mathbf{A}$ 是对称负定的; 对于某些 i , 当 $\lambda_i = 0$ 时, $\mathbf{A}$ 是奇异的。

⑥ 收敛速度是衡量一个最优化计算方法有效性的指标^[1], 下面对其进行讨论。

设一个优化算法产生的迭代点列 $\{\mathbf{x}_k\}$ 在某种范数意义下收敛到最优点 $\mathbf{x}^*$, 即:

$$\lim_{k \rightarrow \infty} \|\mathbf{x}_k - \mathbf{x}^*\| = 0 \quad (1-8)$$

可根据连续两次迭代误差的商来定义该优化算法的收敛速度。如果满足如下条件, 则称该优化算法具有 Q 线性收敛速度 (Q -linear rate), Q 代表商 (quotient)。对于所有充分大的 k , 存在一个常数 $r \in (0, 1)$, 满足:

$$\frac{\|\mathbf{x}_{k+1} - \mathbf{x}^*\|}{\|\mathbf{x}_k - \mathbf{x}^*\|} \leq r \quad (1-9)$$

式(1-9)表示在每次迭代中,迭代点到最优点 x^* 的距离至少按照某一个恒定的因子减小,如序列 $\{1+(0.5)^k\}$ 线性地收敛到 1,此处 $r=0.5$ 。

如果满足如下条件,则称该优化算法具有 Q 超线性收敛速度(Q -superlinear rate):

$$\lim_{k \rightarrow \infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|} = 0 \quad (1-10)$$

如序列 $\{1+k^{-k}\}$ 超线性地收敛到 1,此处 $r=0$ 。

如果满足如下条件,则称该优化算法具有 Q 二次收敛速度(Q -quadratic rate)。对于所有充分大的 k ,存在一个正常数 M ,满足:

$$\frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|^2} \leq M \quad (1-11)$$

如序列 $\{1+(0.5)^{2^k}\}$ 二次收敛到 1,此处 $M=1$ 。

从式(1-9)和式(1-11)可见,收敛速度取决于 r 和 M 。 r 和 M 值的大小不仅取决于所采用的优化算法,而且也取决于优化问题本身的性质。如果不考虑这两个值的大小,二次收敛序列最终总是快于线性收敛序列。

显然,任何二次收敛的序列也会超线性地收敛;任何超线性收敛的序列也会线性地收敛。我们也可以定义更高一次的收敛速度,如三次和四次收敛速度,但在实际应用中并不常用。一般情况下,我们可以定义 $Q-p$ 次收敛速度,对于所有充分大的 k ,存在一个正常数 M 和一个正常数 $p(p>1)$,满足:

$$\frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|^p} \leq M \quad (1-12)$$

一般认为,具有超线性收敛速度和二次收敛速度的方法是比较快的。但应该注意,对任何一个算法,收敛速度的理论结果不能保证算法在实际执行时就一定有好的实际计算效果。

⑦ 设 $S \in \mathbf{R}^n$,如果任意两点之间的连接线段仍然在该集合内,则 S 为凸集(convex set)^[1,2],如图 1-1 所示,即,对于任意 $x_1, x_2 \in S$,有:

$$\alpha x_1 + (1-\alpha)x_2 \in S, \quad \alpha \in [0,1] \quad (1-13)$$

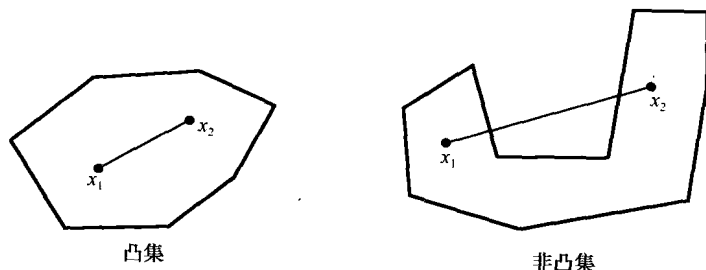


图 1-1 凸集和非凸集的定义