

最优化方法

及其在
机械行业中的应用

| 罗中华 著 |



电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

0224/57

2008

最优化方法及其在机械行业中的应用

罗中华 著

电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

本书共计 9 章和 1 个附录。第 1~6 章, 论述最优化理论与方法; 第 7~9 章, 阐述作者在最优化设计方面的研究实例; 最后, 附录 A 为用 C++ 语言编写的多层压配组合挤压凹模疲劳强度优化设计程序实例。

第 1 章, 阐述二次函数、梯度与 Hessian 矩阵、多元函数的 Taylor 展开、凸集与凸函数、极小点和最优解的充要条件和下降迭代法等最优化设计理论的基本概念; 第 2 章, 论述一维优化方法中的平分法、黄金分割法 (0.618 法) 和抛物线插值法; 第 3 章, 论述无约束最优化问题求解方法中的梯度法、共轭梯度法、牛顿法、变尺度法、模式搜索法、方向加速法 (Powell 法) 和单纯形法; 第 4 章, 论述约束最优化问题求解方法中的三种罚函数法、复合形法、半惩罚函数法和增广拉格朗日乘子法; 第 5 章, 讨论多目标函数的优化设计方法中的理想点法、线性加权法、乘除法和极大极小法; 第 6 章, 简介遗传算法; 第 7 章, 论述两个最优化方法在压力加工中的应用实例; 第 8 章, 论述七个最优化方法在机构设计中的应用实例; 第 9 章, 论述两个最优化方法在拟合公式中的应用实例。

本书可作为高等院校机械类本科生和研究生的教材, 对设计工程师和研究学者也颇有参考价值。

未经许可, 不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有, 侵权必究。

图书在版编目 (CIP) 数据

最优化方法及其在机械行业中的应用 / 罗中华著. —北京: 电子工业出版社, 2008.2

ISBN 978-7-121-05337-5

I. 最… II. 罗… III. 最佳化—数学方法—应用—机械工业 IV. 0224 TH

中国版本图书馆 CIP 数据核字 (2008) 第 003908 号

策划编辑: 田领红

责任编辑: 宋兆武 陈心中

印 刷: 北京市李史山胶印厂

装 订:

出版发行: 电子工业出版社

北京市海淀区万寿路 173 信箱 邮编 100036

开 本: 787×1092 1/16 印张: 14.75 字数: 369 千字

印 次: 2008 年 2 月第 1 次印刷

印 数: 4 000 册 定价: 25.00 元

凡所购买电子工业出版社图书有缺损问题, 请向购买书店调换。若书店售缺, 请与本社发行部联系, 联系及邮购电话: (010) 88254888。

质量投诉请发邮件至 zlt@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线: (010) 88258888。

前 言

最优化问题可追溯到公元前 300 年左右古希腊著名的数学家欧几里得 (Euclid), 他指出在周长相同的一切矩形中, 以正方形的面积为最大。17 世纪, 牛顿、莱布尼兹等著名数学家创立了微积分, 为最优化方法的研究提供了理论基础。

机械优化设计是计算机应用技术、最优化原理与方法 and 机械设计相结合的一门先进设计技术。它是最优化理论和方法为基础, 以计算机为运算工具来寻求最优的机械设计参数的一种现代设计方法。

由于经济全球化, 产品的竞争日趋激烈, 对产品设计质量要求越来越高、设计周期要求越来越短, 传统设计已经不能适应工业发展的需要。为了使机械设计师较好地掌握优化设计方法和将优化设计方法应用于机械设计中, 作者在 20 多年从事最优化设计研究的基础上撰写了本书。

20 多年来, 作者对最优化方法进行了不懈的研究, 取得了一些研究成果。

在最优化理论和方法方面的研究成果有: ①用数学中范数概念正确地证明了 n 维直交定理; ②提出了一种求解最优化问题可行解的快速有效方法; ③提出了半惩罚函数最优化方法。

在最优化设计方面的研究成果有: ①多层压配组合挤压凹模优化设计; ②冲裁零件排样优化设计; ③飞剪机剪切机构的优化设计; ④蟹爪式装载机扒取机构的优化设计; ⑤热锯机送进机构的优化设计; ⑥ J88 型冷挤压机拉力肘杆机构优化设计; ⑦八连杆双动拉延压力机内外滑块执行机构优化设计; ⑧双动拉延压力机内滑块十杆机构的优化设计; ⑨有心扇形滑移线场的近似解析式的拟合; ⑩假塑性塑料熔体在鱼尾形流道中的近似流动公式的拟合等。

这些研究成果中, 有些是作者首次提出来的, 有些是在其他学者的研究基础上再研究的成果。在此, 作者向这些研究学者表示崇高的敬意。

本书的出版得到了上海建桥学院和机电工程系的资助。在撰写过程中, 实训中心的部分教师承担了部分打印工作。在此, 向他们表示衷心的感谢。

由于作者水平有限, 书中错误和不妥之处一定不少, 恳请各位专家、学者批评指正。

罗中华
2007 年 11 月

目 录

第 1 章 最优化问题基础知识	1
1.1 最优化问题的实例	1
1.1.1 利润问题	1
1.1.2 双臂架的最优化设计	2
1.2 最优化问题的基本概念	3
1.2.1 最优化问题的一般形式	3
1.2.2 最优化问题的分类	5
1.3 有关线性代数方面的知识	5
1.3.1 正定矩阵	5
1.3.2 Euclid 空间	6
1.3.3 向量的范数与 Cauchy-Schwarz 不等式	6
1.4 二次函数	8
1.5 梯度与 Hessian 矩阵	9
1.5.1 梯度	9
1.5.2 Hessian 矩阵	11
1.6 多元函数的 Taylor 展开	11
1.7 凸集与凸函数	12
1.7.1 凸集	12
1.7.2 凸函数	13
1.8 极小点与最优解的充要条件	16
1.8.1 极小点的定义	16
1.8.2 最优解的充要条件	18
1.9 下降迭代法及其收敛性	22
1.9.1 下降迭代法	22
1.9.2 收敛速度与收敛准则	24
第 2 章 一维搜索	25
2.1 搜索区间的确定	25
2.1.1 单谷函数与搜索区间	25
2.1.2 确定搜索区间的进退法	26
2.2 平分法	30
2.2.1 利用函数导数的平分法	30
2.2.2 仅利用函数本身的平分法	32
2.3 黄金分割法	34
2.4 抛物线插值法	37
第 3 章 无约束最优化问题的求解方法	40
3.1 梯度法	40

3.2	共轭梯度法	43
3.3	牛顿法	50
3.4	变尺度法	52
3.4.1	变尺度法的基本思想	52
3.4.2	一族变尺度法的形成	54
3.4.3	DFP 变尺度法	56
3.4.4	BFGS 变尺度法	63
3.5	模式搜索法	63
3.6	方向加速法	66
3.7	单纯形法	78
第 4 章	约束最优化问题的求解方法	82
4.1	罚函数法	82
4.1.1	外点罚函数法	83
4.1.2	内点罚函数法	87
4.1.3	扩展的内点罚函数法	90
4.2	复合形法	91
4.3	半惩罚函数法	94
4.3.1	半惩罚函数法提出的原因	94
4.3.2	半惩罚函数法	94
4.4	增广拉格朗日乘子法	96
4.4.1	等式约束最优化问题	96
4.4.2	不等式约束最优化问题	98
4.4.3	一般约束最优化问题	99
第 5 章	多目标函数的优化设计方法	101
5.1	引言	101
5.2	多目标函数优化设计的求解方法	101
5.2.1	理想点法	101
5.2.2	线性加权法	102
5.2.3	乘除法	102
5.2.4	极大极小法	103
第 6 章	遗传算法简介	104
6.1	引言	104
6.2	遗传算法及其特点	104
6.2.1	遗传算法的基本思路及其特点	104
6.2.2	遗传算法的基本步骤	105
6.3	遗传算法的实现方法	106
6.3.1	编码方法	106
6.3.2	适应度函数	109
6.3.3	遗传算子	111
6.3.4	遗传算法运行参数的选择	115

6.4	遗传算法应用实例	116
第7章	最优化方法在压力加工中的应用	119
7.1	多层压配组合挤压凹模优化设计	119
7.1.1	组合凹模的特点	119
7.1.2	整体凹模受力分析	120
7.1.3	组合凹模的力学分析	122
7.1.4	组合凹模的强度校核	124
7.1.5	多层压配组合挤压成形凹模的优化设计	127
7.2	冲裁零件排样优化设计	132
7.2.1	冲裁零件排样	132
7.2.2	冲裁零件排样优化目标函数的建立	134
7.2.3	条料宽度和排样单元步长的计算	135
7.2.4	隔离开关 GN2-10/2000 的支承条零件排样优化设计	138
第8章	最优化方法在机构设计中的应用	139
8.1	飞剪机剪切机构的优化设计	139
8.1.1	对飞剪机剪切机构的要求	140
8.1.2	飞剪机构的运动分析	140
8.1.3	剪切机构的优化设计数学模型的建立	143
8.1.4	50t 飞剪机剪切机构的优化设计	145
8.2	蟹爪式装载机扒取机构的优化设计	147
8.2.1	扒取机构	147
8.2.2	扒取机构的运动学分析	148
8.2.3	扒取机构优化设计的数学模型的建立	149
8.2.4	矿用蟹爪式装载机的扒取机构优化设计实例	153
8.3	热锯机送进机构的优化设计	156
8.3.1	热锯机四连杆送进机构	156
8.3.2	送进机构运动学分析	157
8.3.3	送进机构优化设计数学模型的建立	158
8.3.4	热锯机送进机构优化设计实例	160
8.4	J88 型冷挤压机拉力肘杆机构优化设计	161
8.4.1	冷挤压机概述	161
8.4.2	J88 型冷挤压机拉力肘杆机构	162
8.4.3	拉力肘杆机构非独立构件尺寸参数的计算	162
8.4.4	拉力肘杆机构的运动分析	164
8.4.5	拉力肘杆机构优化设计数学模型的建立	166
8.4.6	J88-400 型冷挤压机拉力肘杆机构优化设计	168
8.5	JA88 型冷挤压机压力肘杆机构优化设计	169
8.5.1	JA88 型冷挤压机压力肘杆机构	169
8.5.2	压力肘杆机构非独立构件尺寸参数的计算	170
8.5.3	压力肘杆机构的运动分析	171

8.5.4	压力肘杆机构的受力分析与曲柄轴摩擦扭矩的计算	172
8.5.5	压力肘杆机构优化设计数学模型的建立	175
8.5.6	JA88-160 型冷挤压机压力肘杆机构优化设计	177
8.6	八连杆双动拉延压力机内外滑块执行机构优化设计	177
8.6.1	板料拉延工艺	177
8.6.2	八连杆双动拉延压力机内外滑块执行机构	179
8.6.3	机构构件非独立构件尺寸参数的计算	179
8.6.4	内、外滑块位移和速度公式的推导	182
8.6.5	八连杆内、外滑块执行机构优化设计目标函数的建立	184
8.6.6	八连杆双动拉延压力机内外滑块执行机构优化设计实例	187
8.7	双动拉延压力机内滑块十杆机构的优化设计	189
8.7.1	内滑块十杆机构	189
8.7.2	非独立机构构件尺寸参数的计算	190
8.7.3	内滑块位移和速度公式的推导	191
8.7.4	内滑块十杆机构优化设计数学模型的建立	193
8.7.5	用半惩罚函数法求内滑块十杆机构优化设计的最优解	197
8.7.6	双动拉延压力机内滑块十杆机构的优化设计实例	197
第 9 章	最优化方法在拟合公式中的应用	199
9.1	有心扇形滑移线场的近似解析式的拟合	199
9.1.1	有心扇形滑移线场的近似解析式	199
9.1.2	有心扇形场的差分公式	202
9.1.3	用优化方法确定有心扇形滑移近似解析式	202
9.1.4	有心扇形滑移线场的近似解析式的应用	203
9.2	假塑性塑料熔体在鱼尾形流道中的近似流动公式的拟合 ^[53]	204
9.2.1	假塑性塑料熔体在鱼尾形流道中的流动公式的推导	204
9.2.2	用优化方法确定插值公式	207
附录 A	多层压配组合挤压凹模疲劳强度优化设计 C++ 程序	211
	参考文献	223

第 1 章 最优化问题基础知识

1.1 最优化问题的实例

1.1.1 利润问题

某煤厂生产市场上畅销的褐煤和烟煤，两种煤每生产 1 吨的利润分别为 4 元和 3 元。该煤厂的破碎、过筛和清洗分厂每天工作时间分别不超过 12 小时、10 小时和 8 小时，生产 1 吨褐煤需要 3 小时破碎、3 小时过筛和 4 小时清洗；生产 1 吨烟煤需要 4 小时破碎、3 小时过筛和 2 小时清洗。问该厂每天生产褐煤和烟煤各多少吨才能达到最大的利润。

设每天生产褐煤 x_1 吨、烟煤 x_2 吨，则每天的利润为

$$f(x_1, x_2) = 4x_1 + 3x_2$$

由于受该厂生产条件的限制，必须满足下列约束条件：

(1) 破碎工作时间不超过 12 小时，即

$$3x_1 + 4x_2 \leq 12$$

(2) 过筛工作时间不超过 10 小时，即

$$3x_1 + 3x_2 \leq 10$$

(3) 清洗工作时间不超过 8 小时，即

$$4x_1 + 2x_2 \leq 8$$

(4) x_1 和 x_2 非负，即

$$x_1, x_2 \geq 0$$

综上所述，该利润问题的最优化数学模型为

$$\max \{4x_1 + 3x_2\} \tag{1.1.1}$$

满足于

$$\begin{aligned} 3x_1 + 4x_2 &\leq 12 \\ 3x_1 + 3x_2 &\leq 10 \\ 4x_1 + 2x_2 &\leq 8 \\ x_1, x_2 &\geq 0 \end{aligned} \tag{1.1.2}$$

1.1.2 双臂架的最优化设计

考虑由空心圆管构成的对称双臂架，如图 1.1.1 所示。已知顶点承受的载荷为 $2P$ 、两支柱的跨距为 $2L$ 、圆管材料的弹性模量为 E 、材料的泊松系数为 ν 、屈服强度 σ_y 和密度为 ρ 。如何选择圆管的平均直径、壁厚和双臂架的高度，使双臂架的重量最轻（即最节省材料）。

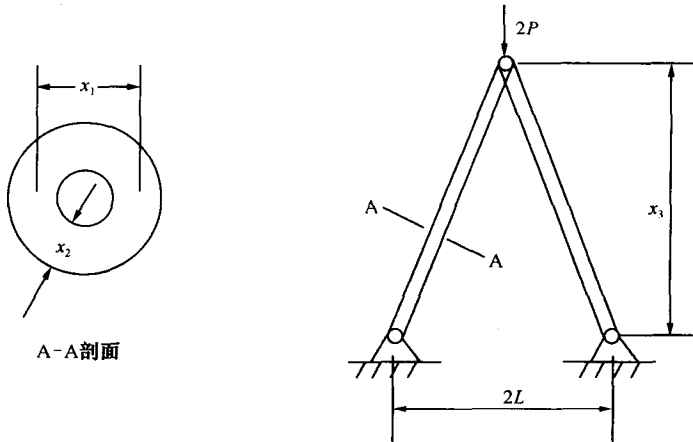


图 1.1.1 双臂架

设双臂架所使用的圆管的平均直径为 x_1 、壁厚为 x_2 和双臂架的高度 x_3 ，则双臂架的重量 W 为

$$W = 2\pi\rho x_1 x_2 \sqrt{L^2 + x_3^2}$$

必须满足以下约束条件：

(1) 圆管中的应力不超过材料的屈服强度 σ_y ，即

$$P\sqrt{L^2 + x_3^2} \leq \pi\sigma_y x_1 x_2 x_3$$

(2) 保证在载荷作用下圆管不发生失稳，由材料力学可知圆管的临界载荷为 $P_{cr} = \pi^2 EI / (L^2 + x_3^2)$ ，其中，截面惯性矩为 $I = \pi x_1 x_2 (x_1^2 + x_2^2) / 8$ 。整理后，可得

$$8P(L^2 + x_3^2)^{3/2} \leq \pi^3 E x_1 x_2 x_3 (x_1^2 + x_2^2)$$

(3) 由于空间的限制，双臂架的高度 x_3 不超过某一极限高度 h_{\max} ，即

$$x_3 \leq h_{\max}$$

(4) 圆管成立的条件为

$$x_1 - x_2 \geq d_{\min}$$

式中， d_{\min} 为制造圆管所允许的最小内径。

(5) 由于制造上的原因，圆管的平均直径与壁厚之比不得超过某一常数 c ，即

$$x_1 \leq cx_2$$

(6) 保证薄臂壳不发生局部弯曲或翘曲, 要求圆管应力小于薄臂壳不发生局部弯曲的极限应力为

$$\sigma_s = \frac{2 E x_2}{x_1 \sqrt{3(1-\nu^2)}} \quad (11)$$

则

$$P \sqrt{3(1-\nu^2)} \sqrt{L^2 + x_3^2} \leq 2\pi E x_2^2 x_3$$

(7) x_1 、 x_2 和 x_3 大于零, 即

$$x_1, x_2, x_3 > 0$$

综上所述, 双臂架的最优化设计的数学模型为

$$\min \left\{ 2\pi\rho x_1 x_2 \sqrt{L^2 + x_3^2} \right\} \quad (1.1.3)$$

满足于

$$\begin{aligned} P \sqrt{L^2 + x_3^2} &\leq \pi \sigma_y x_1 x_2 x_3 \\ 8P(L^2 + x_3^2)^{3/2} &\leq \pi^3 E x_1 x_2 x_3 (x_1^2 + x_2^2) \\ x_3 &\leq h_{\max} \\ x_1 - x_2 &\geq d_{\min} \\ x_1 &\leq c x_2 \\ P \sqrt{3(1-\nu^2)} \sqrt{L^2 + x_3^2} &\leq 2\pi E x_2^2 x_3 \\ x_1, x_2, x_3 &> 0 \end{aligned} \quad (1.1.4)$$

1.2 最优化问题的基本概念

1.2.1 最优化问题的一般形式

在研究最优化问题中, 经常将变量 x_1, x_2, \dots, x_n 视为 n 维向量空间 R^n 中的一个向量 X 的 n 个分量, 即

$$X = \{x_1, x_2, \dots, x_n\}^T$$

最优化问题一般可表示为如下标准形式:

$$\begin{aligned} \min f(X) & \quad (X \in R^n) \\ \text{满足于} \quad g_i(X) &\leq 0 \quad (i=1, 2, \dots, m) \\ h_j(X) &= 0 \quad (j=1, 2, \dots, l) \end{aligned} \quad (1.2.1)$$

其他形式的最优化问题可以转化为如上最优化问题的标准形式, 例如, 求 $\max f(X)$ 可转化为求 $\min \{-f(X)\}$, 同理, $g_i(X) \geq 0$ 亦可转化为标准形式 $-g_i(X) \leq 0$ 。

在式 (1.2.1) 中, $f(X)$ 称为目标函数 (或者称为评价函数), 用它来评价一个设计方案好坏的具体的数值评判标准; $g_i(X) \leq 0$ ($i=1, 2, \dots, m$) 称为不等式约束条件, m 为不等式约

束条件的个数; $h_j(\mathbf{X})=0$ ($j=1,2,\dots,l$) 称为等式约束条件, l 为等式约束条件的个数 [注: $l < n$]. 不等式约束条件和等式约束条件统称为约束条件。

满足所有约束条件的向量 $\mathbf{X} \in \mathbf{R}^n$ 称为可行点 (或称为可行解), 所有可行点组成的集合称为可行解集, 可行解集用 D 来表示, 则有

$$D = \{ \mathbf{X}: g_i(\mathbf{X}) \leq 0 \quad (i=1,2,\dots,m); h_j(\mathbf{X})=0 \quad (j=1,2,\dots,l); \mathbf{X} \in \mathbf{R}^n \} \quad (1.2.2)$$

最优化问题的标准形式可表示为

$$\begin{aligned} & \min f(\mathbf{X}) \\ & \mathbf{X} \in D \end{aligned} \quad (1.2.3)$$

最优化问题 (1.2.1) 的求解是指在可行解集 D 中找一点 \mathbf{X}^* , 使

$$f(\mathbf{X}^*) = \min f(\mathbf{X}) \quad \text{对于所有 } \mathbf{X} \in D$$

满足于

$$g_i(\mathbf{X}^*) \leq 0 \quad (i=1,2,\dots,m)$$

$$h_j(\mathbf{X}^*) = 0 \quad (j=1,2,\dots,l)$$

\mathbf{X}^* 称为最优化问题 (1.2.1) 的最优点, $f(\mathbf{X}^*)$ 称为最优值, $\{\mathbf{X}^*, f(\mathbf{X}^*)\}$ 称为最优解。

下面, 我们用利润问题来说明最优化问题的基本概念, 设计变量 $\mathbf{X} \in \mathbf{R}^2$, 则

$$\mathbf{X} = \{x_1, x_2\}^T$$

优化问题 (1.2.1) 中的目标函数

$$f(\mathbf{X}) = 4x_1 + 3x_2$$

约束条件为

$$g_1(\mathbf{X}) = 3x_1 + 4x_2 - 12 \leq 0$$

$$g_2(\mathbf{X}) = 3x_1 + 3x_2 - 10 \leq 0$$

$$g_3(\mathbf{X}) = 4x_1 + 2x_2 - 8 \leq 0$$

$$g_4(\mathbf{X}) = -x_1 \leq 0$$

$$g_5(\mathbf{X}) = -x_2 \leq 0$$

这个优化问题的可行解集为 $D = \{ \{x_1, x_2\}^T: 3x_1 + 4x_2 - 12 \leq 0, \dots, -x_2 \leq 0, \mathbf{X} \in \mathbf{R}^2 \}$ 。

$\mathbf{X} = \{1, 1\}^T \in D$, 就是一个可行点。利润问题的可行解集如图 1.2.1 所示。

在实际最优化问题中, 每个可行点就是一个可行的设计方案。不同的可行点对应着不同的目标函数值, 目标函数值的大小就是衡量一个设计方案的好坏的数量表征, 因此, 目标函数亦称为评价函数。

用图解法或其他方法求得上述利润问题的最优点为 $\mathbf{X}^* = \{0.8, 2.4\}^T$, 最优值为 $f(\mathbf{X}^*) = 10.4$ 。

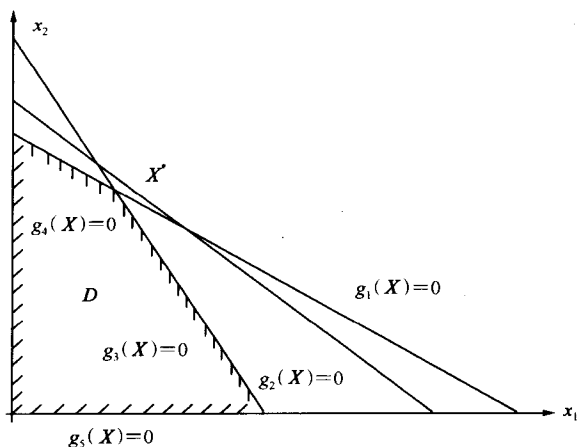


图 1.2.1 利润问题的可行解集

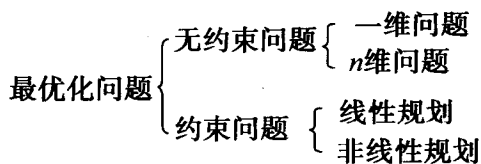
1.2.2 最优化问题的分类

在最优化问题 (1.2.1) 中, 如果 $m=l=0$, 最优化问题转化为没有约束的优化问题, 即

$$\min f(\mathbf{X}) \quad (1.2.4)$$

称为无约束问题; 当 m 和 l 不同时为零时, 称为约束问题。在约束问题中, 如果 $f(\mathbf{X})$ 、 $g_i(\mathbf{X}) \leq 0$ 和 $h_j(\mathbf{X})=0$ 都是线性函数称为线性规划, 否则称为非线性规划。上节中, 利润问题为线性规划, 双臂架最优化问题为非线性规划。在无约束问题中, 如果目标函数是一元函数称为一维问题, 如果目标函数的自变量 $\mathbf{X} \in \mathbf{R}^n$ 称为 n 维问题。

综上所述, 最优化问题的分类可归纳为



1.3 有关线性代数方面的知识

1.3.1 正定矩阵

定义 1.3.1 设 A 为 $n \times n$ 阶对称矩阵, 若对于任意向量 $\mathbf{X} \in \mathbf{R}^n$, 且 $\mathbf{X} \neq \mathbf{0}$ 向量, 都有 $\mathbf{X}^T \mathbf{A} \mathbf{X} > 0$, 则称 A 为正定矩阵; 若对于任意向量 $\mathbf{X} \in \mathbf{R}^n$, 都有 $\mathbf{X}^T \mathbf{A} \mathbf{X} \geq 0$, 则称 A 为半正定矩阵。若对于任意向量 $\mathbf{X} \in \mathbf{R}^n$, 且 $\mathbf{X} \neq \mathbf{0}$ 向量, 都有 $\mathbf{X}^T \mathbf{A} \mathbf{X} < 0$, 则称 A 为负定矩阵; 若对于任意向量 $\mathbf{X} \in \mathbf{R}^n$, 都有 $\mathbf{X}^T \mathbf{A} \mathbf{X} \leq 0$, 则称 A 为半负定矩阵。

定理 1.3.1 (Sylvester 定理) 一个 $n \times n$ 阶对称矩阵 A 为正定矩阵的充要条件是矩阵的

各阶主子式

$$\Delta_1 = a_{11}, \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \dots, \Delta_n = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{12} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

都大于零（定理的证明见文献[2]P337~P339）。

1.3.2 Euclid 空间

定义 1.3.2 设 V 是实数域 R 上的线性空间，如果在 V 上定义一个实数 (X, Y) ，它满足下列条件：

- (1) $(X, Y) = (Y, X)$
- (2) $(X + Y, Z) = (X, Z) + (Y, Z)$
- (3) $(\alpha X, Y) = \alpha(X, Y)$
- (4) $(X, X) \geq 0$ ，当且仅当 $X=0$ 向量时， $(X, X) = 0$

这里 X, Y, Z 是 V 中的任意向量， α 为任意实数，则称这样的实数 (X, Y) 为向量 X 与 Y 的内积，而定义了内积的线性空间 V 称为 Euclid 空间。

在向量空间 R^n 中，可定义内积为

$$(X, Y) = \sum_{i=1}^n x_i y_i$$

式中 $X, Y \in R^n$ ， x_i, y_i 分别为向量 X, Y 的第 i 个元素。显然，上式满足内积定义的条件(1)~(4)，

所以， $(X, Y) = \sum_{i=1}^n x_i y_i$ 可作为向量空间 R^n 的一种内积。

在向量空间 R^n 中，定义内积 $(X, Y) = \sum_{i=1}^n x_i y_i$ ，这时，向量空间 R^n 就成为了 Euclid 空间。

1.3.3 向量的范数与 Cauchy-Schwarz 不等式

定义 1.3.3 设 V 是数域 K 上的一个线性空间， $X \in V$ 为任意向量， X 对应一个非负的实数 $\|X\|$ ， $\|X\|$ 满足如下三个条件：

- (1) $\|X\| \geq 0$ ，当且仅当 $X=0$ 向量时， $\|X\| = 0$
- (2) 对于任意常数 $\alpha \in K$ ， $\|\alpha X\| = |\alpha| \|X\|$
- (3) 对于任意 $X, Y \in V$ ，如下“三角不等式”成立：

$$\|X + Y\| \leq \|X\| + \|Y\| \quad (1.3.1)$$

则称实数 $\|X\|$ 为向量 X 的范数。

在向量空间 R^n 中，可定义向量 X 的范数 $\|X\|$ 为

$$\|X\| = (X, X)^{1/2} = \left(\sum_{i=1}^n x_i^2 \right)^{1/2} \quad (1.3.2)$$

在验证之前, 我们先引进 Cauchy-Schwarz 不等式。

定理 1.3.2 (Cauchy-Schwarz 不等式) 设 $X, Y \in R^n$, 则有

$$|(X, Y)| \leq \|X\| \|Y\| \quad (1.3.3)$$

证 对于任意实数 α , 恒有

$$(\alpha X + Y, \alpha X + Y) \geq 0$$

将上式展开, 得

$$\alpha^2 (X, X) + 2\alpha (X, Y) + (Y, Y) \geq 0$$

上式左边是 α 的二次三项式, 由于它的符号不变, 所以它的判别式

$$(X, Y)^2 - (X, X)(Y, Y) \leq 0$$

由上式可得

$$|(X, Y)| \leq \|X\| \|Y\|$$

证毕。

现在, 来验证由式 (1.3.2) 定义的 $\|X\| = (X, X)^{1/2} = \left(\sum_{i=1}^n x_i^2 \right)^{1/2}$ 是一种范数。显然, 它满足

向量范数的条件 (1) 和条件 (2)。下面, 证明它满足范数的“三角不等式”条件。

利用 Cauchy-Schwarz 不等式, 有

$$\begin{aligned} (X+Y, X+Y) &= (X, X) + 2(X, Y) + (Y, Y) \\ &\leq \|X\|^2 + 2\|X\|\|Y\| + \|Y\|^2 \\ &\leq (\|X\| + \|Y\|)^2 \end{aligned}$$

上式两边开平方, 得

$$\|X+Y\| \leq \|X\| + \|Y\|$$

上式说明 $\|X\| = (X, X)^{1/2} = \left(\sum_{i=1}^n x_i^2 \right)^{1/2}$ 满足范数条件 (3), 因此, 它是向量空间 R^n 中的一种范数。

另外, 还可以通过对称正定矩阵来定义内积和范数。设 A 为 $n \times n$ 阶对称正定矩阵, 在向量空间 R^n 中, 亦可定义内积为

$$(X, Y) = X^T A Y \quad (1.3.4)$$

定义范数为

$$\|X\| = (X^T A X)^{1/2} \quad (1.3.5)$$

式中 $X, Y \in R^n$ 。

不难证明 $(X, Y) = X^T A Y$ 满足内积条件 (1) ~ (4), 因此, $(X, Y) = X^T A Y$ 可作为向量空间 R^n 中的一种内积。

显然, $\|X\| = (X^T A X)^{1/2}$ 满足范数条件 (1) 和条件 (2)。下面, 证明 $\|X\| = (X^T A X)^{1/2}$ 满

足范数的“三角不等式”条件。

由于 A 为 $n \times n$ 阶对称正定矩阵, 对于任意 $X, Y \in R^n$, α 为任意实数, 有

$$(\alpha X + Y)^T A (\alpha X + Y) \geq 0$$

由上式得

$$\alpha^2 X^T A X + 2\alpha X^T A Y + Y^T A Y \geq 0$$

上式左边是 α 的二次三项式, 由于它的符号不变, 所以, 它的判别式

$$(X^T A Y)^2 - X^T A X Y^T A Y \leq 0 \quad (1.3.6)$$

由上式得

$$|X^T A Y| \leq (X^T A X)^{1/2} (Y^T A Y)^{1/2}$$

即

$$\begin{aligned} |(X, Y)| &\leq \|X\| \|Y\| \\ \|X + Y\|^2 &= (X + Y)^T A (X + Y) \\ &= X^T A X + 2X^T A Y + Y^T A Y \\ &\leq \|X\|^2 + 2\|X\| \|Y\| + \|Y\|^2 \end{aligned} \quad (1.3.7)$$

故有

$$\|X + Y\| \leq \|X\| + \|Y\|$$

因此, $\|X\| = (X^T A X)^{1/2}$ 可作为向量空间 R^n 中的一种范数。

1.4 二次函数

在向量空间 R^n 中, 线性函数可以表示为

$$f(x_1, x_2, \dots, x_n) = \sum_{i=1}^n b_i x_i + c \quad (1.4.1)$$

用向量形式表示

$$f(X) = B^T X + c \quad (1.4.2)$$

式中 $B = \{b_1, b_2, \dots, b_n\}^T$, c 为常数。

在最优化研究中, 二次函数起着重要的作用。二次函数的一般形式为

$$f(x_1, x_2, \dots, x_n) = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j + \sum_{i=1}^n b_i x_i + c \quad (1.4.3)$$

式中 a_{ij} , b_i 和 c 为常数。用矩阵向量可表示为

$$f(X) = \frac{1}{2} X^T A X + B^T X + c \quad (1.4.4)$$

其中 $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$, $B = \{b_1, b_2, \dots, b_n\}^T$, c 为常数。

A 为对称正定矩阵的二次函数是我们最关注的函数,这是因为许多最优化方法都是建立在 A 为对称正定矩阵的二次函数的基础上。

例 1.4.1 把二次函数

$$f(x_1, x_2, x_3) = 3x_1^2 + x_2^2 + 2x_3^2 - 3x_1x_2 + x_1x_3 - 4x_1 + 5x_2 + 2$$

转化为矩阵向量的形式,并验证矩阵 A 是否为正定矩阵。

解

$$f(\mathbf{X}) = \frac{1}{2}(6x_1^2 - 3x_1x_2 + x_1x_3 - 3x_2x_1 + 2x_2^2 + 0x_2x_3 + x_3x_1 + 0x_3x_2 + 4x_3^2) - 4x_1 + 5x_2 + 2$$

$$= \frac{1}{2}(6x_1 - 3x_2 + x_3, -3x_1 + 2x_2 + 0x_3, x_1 + 0x_2 + 4x_3) \begin{Bmatrix} x_1 \\ x_2 \\ x_3 \end{Bmatrix} + (-4, 5, 0) \begin{Bmatrix} x_1 \\ x_2 \\ x_3 \end{Bmatrix} + 2$$

$$= \frac{1}{2}(x_1, x_2, x_3) \begin{bmatrix} 6 & -3 & 1 \\ -3 & 2 & 0 \\ 1 & 0 & 4 \end{bmatrix} \begin{Bmatrix} x_1 \\ x_2 \\ x_3 \end{Bmatrix} + (-4, 5, 0) \begin{Bmatrix} x_1 \\ x_2 \\ x_3 \end{Bmatrix} + 2$$

$$= \frac{1}{2} \mathbf{X}^T \mathbf{A} \mathbf{X} + \mathbf{B}^T \mathbf{X} + c$$

式中 $\mathbf{A} = \begin{bmatrix} 6 & -3 & 1 \\ -3 & 2 & 0 \\ 1 & 0 & 4 \end{bmatrix}$, $\mathbf{B} = \begin{pmatrix} -4 \\ 5 \\ 0 \end{pmatrix}$, $\mathbf{X} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, $c = 2$ 。

A 的各阶主子式分别为

$$A_1 = 6 > 0, \quad A_2 = \begin{vmatrix} 6 & -3 \\ -3 & 2 \end{vmatrix} = 3 > 0, \quad A_3 = \begin{vmatrix} 6 & -3 & 1 \\ -3 & 2 & 0 \\ 1 & 0 & 4 \end{vmatrix} = 10 > 0$$

根据定理 1.3.1 知,矩阵 A 为对称正定矩阵。

1.5 梯度与 Hessian 矩阵

1.5.1 梯度

定义 1.5.1 如果函数 $f(\mathbf{X})$ 的所有一阶偏导数存在,则称

$$\nabla f(\mathbf{X}) = \left(\frac{\partial f(\mathbf{X})}{\partial x_1}, \frac{\partial f(\mathbf{X})}{\partial x_2}, \dots, \frac{\partial f(\mathbf{X})}{\partial x_n} \right)^T \quad (1.5.1)$$