高等学校工科类

研究生教学用书

最优控制理论与应用

Optimal Control Theory and Application

张洪钺　王　青　编著

高等教育出版社

高等学校工科类

研究生教学用书

最优控制理论与应用

Optimal Control Theory and Application

张洪钺 王 青 编著

高等教育出版社

内容提要

本书系统地介绍了最优控制的理论基础、最优控制的计算方法和最优控制的应用。

全书共分12章,第1章绪论,第2章静态最优控制,第3章用变分法解最优控制、第4章极小值原理及其应用,第5章线性二次性型指标的最优控制,第6章动态规划,第7章最优控制的计算方法,第8章随机系统的最优控制,第9章奇异最优控制,第10章对策论与极大极小控制,第11章鲁棒与最优控制,第12章用MATLAB解最优控制问题及应用实例。

本书可作为控制及相关专业的研究生教材和高年级本科生的选修教材,也可供从事相关专业的科研人员和工程技术人员参考。

图书在版编目(CIP)数据

最优控制理论与应用/张洪钺,王青编著.—北京:高等教育出版社,2006.1

ISBN 7-04-018400-1

Ⅰ.最... Ⅱ.①张...②王... Ⅲ.最优控制-数学理论 Ⅳ.0232

中国版本图书馆CIP数据核字(2005)第147945号

策划编辑	刘　英	**责任编辑**	陈思宇	**封面设计**	李卫青
责任绘图	朱　静	**版式设计**	史新薇	**责任校对**	俞声佳
责任印制	孔　源				

出版发行	高等教育出版社	**购书热线**	010-58581118
社　　址	北京市西城区德外大街4号	**免费咨询**	800-810-0598
邮政编码	100011	**网　　址**	http://www.hep.edu.cn
总　　机	010-58581000		http://www.hep.com.cn
		网上订购	http://www.landraco.com
经　　销	蓝色畅想图书发行有限公司		http://www.landraco.com.cn
印　　刷	北京市南方印刷厂	**畅想教育**	http://www.widedu.com
开　　本	787×960　1/16	**版　　次**	2006年1月第1版
印　　张	16	**印　　次**	2006年1月第1次印刷
字　　数	260 000	**定　　价**	28.90元

物料号　18400-00

前　　言

最优控制理论已经有很长的发展历史，但是近些年来最优控制理论及其在实际中的应用又有了新的发展。最优控制在生产、国防乃至经济管理等应用领域，继续发挥着重要作用。

本书在阐明最优控制问题应用背景的基础上，循序渐进地描述了最优控制基本理论，系统地介绍了最优控制系统的设计和计算方法，并以具有实际背景的例子说明理论方法的应用。考虑到最优控制理论及其在应用中的发展，本书增加了鲁棒 H_∞ 最优控制、对策论与极大极小控制等内容，并介绍了控制系统计算机辅助设计与仿真软件 MATLAB 工具在最优控制中的应用，给出了应用实例。

本书在选材范围和数学工具的使用上充分考虑了教学需求和多数学生的知识结构，面向工科学生，注重基本理论的学习和主要方法的掌握，避免过于繁琐的大篇幅数学论证，给出了许多易于学生理解的例子，注重分析和解决实际问题能力的培养，使学生能够应用最优控制理论和方法来解决一些实际的控制问题。

全书共分 12 章，第 1 章绪论，第 2 章静态优化，第 3 章用变分法解最优控制，第 4 章极小值原理及其应用，第 5 章线性二次型指标的最优控制，第 6 章动态规划，第 7 章最优控制的计算方法，第 8 章随机线性系统的最优控制，第 9 章奇异最优控制，第 10 章对策论与极大极小控制，第 11 章鲁棒与最优控制，第 12 章用 MATLAB 解最优控制问题及应用实例。

本书第 1 至 8 章和第 10 章由张洪钺教授编写，第 9、11 和 12 章由王青教授编写。对全书内容的取舍，两位作者进行过深入讨论。在书稿完成之际，对参加本书编写、校对的祝世虎博士生、陈宇博士生、余杨博士生表示衷心的感谢。

限于我们的水平以及时间仓促，虽经多次校对，但书中不妥和错误之处仍在所难免，恳请广大读者予以批评指正。

编　　者

2005 年 9 月于北航

目　　录

第 1 章　绪　　论

在生产过程、军事行动、经济活动以及人类的其他有目的的活动中，常需要对被控系统或被控过程施加某种控制作用以使某个性能指标达到最优，这种控制作用称为最优控制。

最优控制理论已有很长的发展历史。早在上世纪 50 年代初期布绍（Bushaw）就研究了伺服系统的时间最优控制问题，他用几何方法证明了继电式的控制可以用最短的时间将伺服系统的误差调节到 0。以后，拉塞尔（LaSalle）发展了时间最优控制的理论，即所谓 Bang - Bang 控制理论（“Bang - Bang”模拟了继电器动作时所发出的声音）。50 年代空间技术开始获得迅猛的发展。导弹、卫星等都是复杂的多输入 - 多输出的非线性系统，而且在性能上有严格的要求，如消耗燃料要少，飞行速度要快，在重量和可靠性方面也有严格的要求。这种工程上的要求刺激了最优控制理论的发展。人们发现，最优控制问题就其本质来说，是一个变分学的问题。然而，经典变分学只能解决控制作用不受限制的情况。实际上常常碰到控制作用受到限制的情况，这就要求人们开辟求解最优控制的新途径。1953 至 1957 年间美国学者贝尔曼（Bellman）创立了“动态规划”理论，发展了变分学中的哈密顿 - 雅可比（Hamilton - Jacobi）理论。1956 至 1958 年间前苏联学者庞特里亚金（Pontryagin）等创立了“极大值原理”。这两种方法成为了目前最优控制理论的两个柱石。时至今日，最优控制理论的研究无论在深度上和广度上都有了很大的发展，例如，发展了对分布参数系统、随机系统、大系统的最优控制理论的研究，等等。电子计算机的发展为最优控制的实现提供了强有力的手段。毫不夸张地说，最优控制理论仍是一个活跃的科学研究的领域，它在国民经济和国防事业中将继续发挥重要的作用。下面列举几个简单的最优控制的例子。

例 1 - 1　火车快速运行问题。设有一列火车从甲地出发，要求算出容许的控制，使其到达乙地的时间最短。

火车的运动方程

$$m\ddot{x} = u(t) \tag{1 - 1}$$

式中，m 是火车的质量；$\ddot{x}$ 是火车的加速度，为使旅客舒适，其值有限制；$u(t)$ 是产生加速度的控制作用（即推力），其值也应有限制，设

$$|u(t)| \leqslant M \tag{1-2}$$

初始条件

$$x(t_0) = x_0 \qquad \dot{x}(t_0) = 0 \tag{1-3}$$

终端条件

$$x(t_f) = x_f \qquad \dot{x}(t_f) = 0 \tag{1-4}$$

性能指标

$$J(u) = \int_{t_0}^{t_f} \mathrm{d}t = t_f - t_0 \tag{1-5}$$

选择 $u(t)$ 使 $J(u)$ 为最小。

例 1-2 月球软着陆问题。为了使宇宙飞船在月球表面上实现软着陆(即着陆时速度要为0),要寻求着陆过程中发动机推力的最优控制规律,使得燃料的消耗最少。设飞船的质量为 $m(t)$,离月球表面的高度为 $h(t)$,飞船的垂直速度为 $v(t)$,发动机推力为 $u(t)$,月球表面的重力加速度为 g,设不带燃料的飞船质量为 M,初始燃料的质量为 F,则飞船的运动方程可表示为(参见图 1-1)

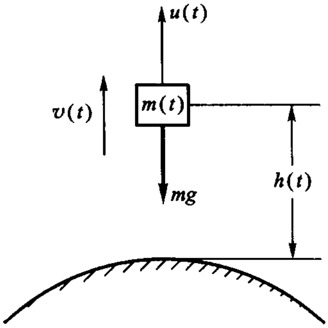

图 1-1 月球软着陆最优控制问题

$$\begin{aligned} \dot{h}(t) &= v(t) \\ \dot{v}(t) &= -g + \frac{u(t)}{m(t)} \\ \dot{m}(t) &= -ku(t) \end{aligned} \tag{1-6}$$

式中 k 为比例系数,表示了推力与燃料消耗率的关系。

初始条件

$$h(t_0) = h_0 \quad v(t_0) = v_0 \quad m(t_0) = M + F \tag{1-7}$$

终端条件

$$h(t_f)=0 \qquad v(t_f)=0 \tag{1-8}$$

容许控制

$$0 \leqslant u(t) \leqslant a \tag{1-9}$$

控制目的是使燃料消耗量最小，即飞船在着陆时的质量保持最大，即

$$J(u)=m(t_f) \tag{1-10}$$

为最大。

例 1－3　生产计划问题。设 $x(t)$ 表示商品存货量，$r(t)\geqslant 0$ 表示对商品的需求率，是已知函数，$u(t)$ 表示生产率，它将由计划人员来选取，故是控制变量。$x(t)$ 满足下面的微分方程

$$\dot{x}(t)=-r(t)+u(t) \qquad t\in[0,t_f] \tag{1-11}$$

$$x(0)=x_0$$

x_0 是初始时刻的商品存货量，且 $x_0\geqslant 0$。从 $x(t)$ 的实际意义来看，显然必须选取生产率 $u(t)$ 使得

$$x(t)\geqslant 0 \qquad t\in[0,t_f] \tag{1-12}$$

其次，生产能力应该有限制，即容许控制为

$$0\leqslant u(t)\leqslant A \qquad t\in[0,t_f] \tag{1-13}$$

这里 $A>0$ 表示最大生产率，另外为了保证满足需求，必须有

$$A>r(t) \qquad t\in[0,t_f] \tag{1-14}$$

假定每单位时间的生产成本是生产率 $u(t)$ 的函数，即 $h[u(t)]$。设 $b>0$ 是单位时间储存单位商品的费用，于是，单位时间的总成本为

$$f[x(t),u(t),t]=h[u(t)]+bx(t) \tag{1-15}$$

由 $t=0$ 到 $t=t_f$ 的总成本为

$$J(u)=\int_0^{t_f} f[x(t),u(t),t]\,\mathrm{d}t \tag{1-16}$$

要求寻找最优控制 $u^*(t)$，使总成本 J 最小。

由上面的例子可见，求解最优控制问题时要给定系统的状态方程、状态变量所满足的初始条件和终端条件、性能指标的形式（时间最短、消耗燃料最少、误差平方积分最小等）以及控制作用的容许范围等。

用数学语言来比较详细地表达最优控制问题的内容：

（1）建立被控系统的状态方程

$$\dot{\boldsymbol{X}}=\boldsymbol{f}[\boldsymbol{X}(t),\boldsymbol{U}(t),t] \tag{1-17}$$

其中，$X(t)$为 n 维状态向量，$U(t)$为 m 维控制向量，$f[X(t),U(t),t]$为 n 维向量函数，它可以是非线性时变向量函数，也可以是线性定常的向量函数。状态方程必须精确地知道。

（2）确定状态方程的边界条件。一个动态过程对应于 n 维状态空间中从一个状态到另一个状态的转移，也就是状态空间中的一条轨线。在最优控制中初态通常是已知的，即

$$X(t_0) = X_0 \tag{1-18}$$

而到达终端的时刻 t_f 和状态 $X(t_f)$则因问题而异。例如，在流水线生产过程中，t_f 是固定的；在飞机快速爬高时，只规定爬高的高度 $X(t_f) = X_f$，而 t_f 是自由的，要求 $t_f - t_0$ 越小越好。终端状态 $X(t_f)$一般属于一个目标集 S，即

$$X(t_f) \in S \tag{1-19}$$

当终端状态是固定的，即 $X(t_f) = X_f$ 时，则目标集退化为 n 维状态空间中的一个点。而当终态满足某些约束条件，即

$$G[X(t_f),t_f] = 0 \tag{1-20}$$

这时 $X(t_f)$处在 n 维状态空间中某个超曲面上。若终态不受约束，则目标集便扩展到整个 n 维空间，或称终端状态自由。

（3）选定性能指标 J。性能指标一般有下面的形式：

$$J = \phi[X(t_f),t_f] + \int_{t_0}^{t_f} L[X(t),U(t),t]\mathrm{d}t \tag{1-21}$$

上述性能指标包括两个部分，即积分指标 $\int_{t_0}^{t_f} L[X(t),U(t),t]\mathrm{d}t$ 和终端指标 $\phi[X(t_f),t_f]$，这种综合性指标所对应的最优控制问题称为波尔扎（Bolza）问题。当只有终端指标时，称为迈耶尔（Mayer）问题；当只有积分指标时，称为拉格朗日（Lagrange）问题。性能指标的确定因问题的性质而异。在导弹截击目标的问题中，要求弹着点的散布度最小，这时可用终端指标来表示。在快速控制问题中，要求系统从一个状态过渡到另一个状态的时间最短，即$\int_{t_0}^{t_f}\mathrm{d}t \to \min$，这就是积分指标。性能指标 J 是控制作用 $U(t)$ 的函数，也就是函数 $U(t)$ 的函数，这种以函数为自变量的函数称为泛函，所以 J 又称为性能泛函。有的文献中也把性能指标称为代价函数、目标函数等。

（4）确定控制作用的容许范围 Ω，即

$$U(t) \in \Omega \tag{1-22}$$

Ω 是 m 维控制空间 R^m 中的一个集合。例如，控制飞机的舵偏角是受限制的，控制电机的电流是受限制的，即有 $|U(t)| \leqslant M$。这时控制作用属于一个闭集。当 $U(t)$ 不受任何限制时，称它属于一个开集。下面将看到处理这两类问题的方法是不同的。Ω 可称为容许集合，属于 Ω 的控制则称为容许控制。

(5) 按一定的方法计算出容许控制 $U(t)$ ($U(t) \in \Omega$)，将它施加于用状态方程描述的系统，使状态从初态 $X(t_0)$ 转移到目标集 S 中的某一个终态 $X(t_f)$，并使性能指标达到最大或最小，即达到某种意义下的最优。

本书将介绍求解最优控制问题的方法：经典变分法，极大(小)值原理，动态规划法，线性二次型最优控制(系统为线性，指标为状态和控制的二次型)，线性二次型高斯控制(系统为线性且有高斯噪声，指标为二次型)，奇异最优控制，微分对策控制(系统受双方控制)，H_∞ 最优鲁棒控制，等等。本书还将介绍最优控制的一些基本的数值求解方法，最后介绍一些 MATLAB 在求解最优控制问题中的应用实例。

第 2 章 静态优化
——函数的极值问题

正如在第 1 章中所述，性能指标 J 是控制作用（注意，它是时间的函数）的函数，即函数的函数，数学上将其定义为泛函。寻求控制作用使性能指标 J 达到极大或极小的问题，就是求泛函极值的问题。求函数的极值要用到函数的微分，而求泛函极值时则要用到泛函的变分。泛函的变分和函数的微分是十分类似的。所以，下面先来复习函数的极值理论，然后再来讨论泛函的极值问题，这样会比较容易理解。函数极值理论可应用于静态最优问题，即使对动态最优问题，若将控制作用参数化，也可将性能泛函的极值问题转化为函数的极值问题。

2.1 无约束条件的函数极值问题

一元函数 $f(x)$ 在 $x = x^*$ 处取极值的必要条件为

$$f'(x^*) = \frac{\mathrm{d}f(x)}{\mathrm{d}x}\bigg|_{x=x^*} = 0 \tag{2-1}$$

当

$$f''(x^*) = \frac{\mathrm{d}^2 f(x)}{\mathrm{d}x^2}\bigg|_{x=x^*} > 0 \tag{2-2}$$

时，$f(x^*)$ 为极小。当

$$f''(x^*) = \frac{\mathrm{d}^2 f(x)}{\mathrm{d}x^2}\bigg|_{x=x^*} < 0 \tag{2-3}$$

时，$f(x^*)$ 为极大。

为简单起见，以下将只讨论极小，式(2-1)和(2-2)一起构成 $f(x^*)$ 为极小值的充分条件。当 $f''(x^*) = 0$ 时，也可能有极小值，不过要检验高阶导数。上述情况可用图 2-1 来表示。R 点是局部极小点，又是总体极小点，U 只是局部极小点，T 是局部极大点，S 是拐点，不是极值点。

例 2-1 求使

$$f(x) = (x - a_1)^2 + (x - a_2)^2 + \cdots + (x - a_n)^2$$

最小的 x。

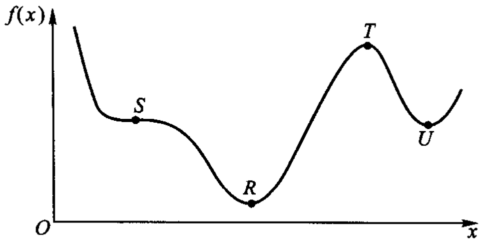

图 2-1 函数的极值点和拐点

解 $f'(x) = 2(x-a_1) + 2(x-a_2) + \cdots + 2(x-a_n) = 0$

$$x = \frac{a_1 + a_2 + \cdots + a_n}{n}$$

$$f''(x^*) = 2n > 0$$

故解 x 使 $f(x)$ 达到极小。本例是著名的最小二乘问题。例如，x 是某个未知的物理量，a_i 是测量值。根据这些测量值可决定 x，使误差的平方和为最小。

下面考虑二元函数 $f(x_1, x_2)$ 的极值问题。设 $f(x_1, x_2)$ 在 $\boldsymbol{X}^* = [x_1^*, x_2^*]^{\mathrm{T}}$ 处取得极小值，记 $f(x_1, x_2) = f(\boldsymbol{X})$，这里 $\boldsymbol{X} = [x_1, x_2]^{\mathrm{T}}$（T 表示转置，$\boldsymbol{X}$ 是列向量）。$f(\boldsymbol{X})$ 在 $\boldsymbol{X} = \boldsymbol{X}^*$ 处取得极小值的必要条件和充分条件可如下求得。将 $f(\boldsymbol{X})$ 在 $\boldsymbol{X} = \boldsymbol{X}^*$ 周围展开为泰勒级数

$$\begin{aligned} f(\boldsymbol{X}) &= f(x_1, x_2) = f(x_1^* + \Delta x_1, x_2^* + \Delta x_2) \\ &= f(\boldsymbol{X}^*) + f'_{x_1}(\boldsymbol{X}^*)\Delta x_1 + f'_{x_2}(\boldsymbol{X}^*)\Delta x_2 + f''_{x_1x_1}(\boldsymbol{X}^*)(\Delta x_1)^2 + \\ &\quad 2f''_{x_1x_2}(\boldsymbol{X}^*)\Delta x_1 \Delta x_2 + f''_{x_2x_2}(\boldsymbol{X}^*)(\Delta x_2)^2 + o[(\Delta x_1)^2, (\Delta x_2)^2] \end{aligned} \tag{2-4}$$

式中

$$f'_{x_1}(\boldsymbol{X}^*) = \left.\frac{\partial f(x_1, x_2)}{\partial x_1}\right|_{\substack{x_1 = x_1^* \\ x_2 = x_2^*}}$$

$$f'_{x_2}(\boldsymbol{X}^*) = \left.\frac{\partial f(x_1, x_2)}{\partial x_2}\right|_{\substack{x_1 = x_1^* \\ x_2 = x_2^*}}$$

$$f''_{x_1x_1}(\boldsymbol{X}^*) = \left.\frac{\partial^2 f(x_1, x_2)}{\partial x_1^2}\right|_{\substack{x_1 = x_1^* \\ x_2 = x_2^*}} \qquad f''_{x_1x_2}(\boldsymbol{X}^*) = \left.\frac{\partial^2 f(x_1, x_2)}{\partial x_1 \partial x_2}\right|_{\substack{x_1 = x_1^* \\ x_2 = x_2^*}}$$

$$f''_{x_2x_2}(\boldsymbol{X}^*) = \left.\frac{\partial^2 f(x_1,x_2)}{\partial x_2^2}\right|_{\substack{x_1=x_1^*\\x_2=x_2^*}}$$

$o[(\Delta x_1)^2,(\Delta x_2)^2]$表示高阶无穷小。将式(2-4)用向量矩阵形式表示

$$f(\boldsymbol{X}) = f(x_1,x_2) = f(\boldsymbol{X}^*) + [\Delta x_1 \quad \Delta x_2]\begin{bmatrix} f'_{x_1} \\ f'_{x_2} \end{bmatrix}_{\boldsymbol{X}=\boldsymbol{X}^*} +$$

$$[\Delta x_1 \quad \Delta x_2]\begin{bmatrix} f''_{x_1x_1} & f''_{x_1x_2} \\ f''_{x_1x_2} & f''_{x_2x_2} \end{bmatrix}_{\boldsymbol{X}=\boldsymbol{X}^*}\begin{bmatrix} \Delta x_1 \\ \Delta x_2 \end{bmatrix} + o[(\Delta x_1)^2,(\Delta x_2)^2]$$

$$= f(\boldsymbol{X}^*) + \Delta\boldsymbol{X}^{\mathrm{T}} f'_{\boldsymbol{X}^*} + \Delta\boldsymbol{X}^{\mathrm{T}} f''_{\boldsymbol{X}^*}\Delta\boldsymbol{X} + o[(\Delta x_1)^2,(\Delta x_2)^2] \tag{2-5}$$

式中

$$\boldsymbol{X} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \qquad \Delta\boldsymbol{X} = \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \end{bmatrix} \qquad f'_{\boldsymbol{X}^*} = [f'_{x_1} \quad f'_{x_2}]^{\mathrm{T}}_{\boldsymbol{X}=\boldsymbol{X}^*}$$

$$f''_{\boldsymbol{X}^*} = \begin{bmatrix} f''_{x_1x_1} & f''_{x_1x_2} \\ f''_{x_1x_2} & f''_{x_2x_2} \end{bmatrix}_{\boldsymbol{X}=\boldsymbol{X}^*} \tag{2-6}$$

由式(2-5)可知,$f(\boldsymbol{X}^*) = f(x_1^*, x_2^*)$取极值的必要条件为

$$f'_{\boldsymbol{X}^*} = 0 \tag{2-7}$$

进一步,若

$$\Delta\boldsymbol{X}^{\mathrm{T}} f''_{\boldsymbol{X}^*}\Delta\boldsymbol{X} > 0 \tag{2-8}$$

则这个极值为极小值。由于$\Delta\boldsymbol{X}$是任意的不为0的向量,要使式(2-8)成立,由矩阵理论可知,二阶导数矩阵(又称为Hessian阵)$f''_{\boldsymbol{X}^*}$必须是正定的。正定阵形式上可表示为

$$f''_{\boldsymbol{X}^*} > 0 \tag{2-9}$$

式(2-7)和(2-9)一起构成了$f(\boldsymbol{X})$在$\boldsymbol{X}^* = [x_1^*, x_2^*]^{\mathrm{T}}$处取极小值的充分条件。

上面的结果不难推广到多元函数的极值问题。设n个变量的多元函数为

$$f(\boldsymbol{X}) = f(x_1, x_2, \cdots, x_n)$$

式中

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

则 $f(X)$ 在 $X = X^* = [x_1^*, x_2^*, \cdots, x_n^*]^T$ 处有极小值的必要条件为一阶导数向量等于零向量,即

$$f'_X = [f'_{x_1}, f'_{x_2}, \cdots, f'_{x_n}]^T_{X=X^*} = 0 \qquad (2-10)$$

进一步,若二阶导数矩阵是正定阵,即

$$f''_{X^*} = \begin{bmatrix} f''_{x_1x_1} & f''_{x_1x_2} & \cdots & f''_{x_1x_n} \\ f''_{x_2x_1} & f''_{x_2x_2} & \cdots & f''_{x_2x_n} \\ & \cdots\cdots & & \\ f''_{x_nx_1} & f''_{x_nx_2} & \cdots & f''_{x_nx_n} \end{bmatrix}_{X=X^*} > 0 \qquad (2-11)$$

则这个极值是极小。式(2-10)和(2-11)一起构成了多元函数 $f(x_1, x_2, \cdots, x_n)$ 在 $X^* = [x_1^*, x_2^*, \cdots, x_n^*]^T$ 处取极小值的充分条件。

由式(2-11)可知,f''_{X^*} 是实对称矩阵。判别实对称矩阵是否为正定有两个常用的方法。一是检验 f''_{X^*} 的特征值,若特征值全部为正,则 f''_{X^*} 是正定的。另一是应用西尔维斯特(Sylvester)判据。根据此判据,若 f''_{X^*} 的各阶顺序主子式均大于0,即

$$\det\begin{bmatrix} f''_{x_1x_1} & f''_{x_1x_2} & \cdots & f''_{x_1x_k} \\ f''_{x_2x_1} & f''_{x_2x_2} & \cdots & f''_{x_2x_k} \\ & \cdots\cdots & & \\ f''_{x_kx_1} & f''_{x_kx_2} & \cdots & f''_{x_kx_k} \end{bmatrix}_{X=X^*} > 0 \quad k = 1,2,\cdots,n \qquad (2-12)$$

则 f''_{X^*} 就是正定的。det $\boldsymbol{A}$ 表示 $\boldsymbol{A}$ 阵的行列式的值。

例 2-2 求下面的多元函数的极值点

$$f(x_1, x_2, x_3) = 2x_1^2 + 5x_2^2 + x_3^2 + 2x_2x_3 + 2x_3x_1 - 6x_2 + 3$$

解

$$f'_{x_1} = \frac{\partial f}{\partial x_1} = 4x_1 + 2x_3 = 0$$

$$f'_{x_2} = \frac{\partial f}{\partial x_2} = 10x_2 + 2x_3 - 6 = 0$$

$$f'_{x_3} = \frac{\partial f}{\partial x_3} = 2x_1 + 2x_2 + 2x_3 = 0$$

由上面三个方程,求得可能的极值点为

$$\boldsymbol{X}^* = [x_1^*, x_2^*, x_3^*]^{\mathrm{T}} = [1, 1, -2]^{\mathrm{T}}$$

二阶导数阵为

$$f''_{X^*} = \begin{bmatrix} 4 & 0 & 2 \\ 0 & 10 & 2 \\ 2 & 2 & 2 \end{bmatrix}$$

用西尔维斯特判据来检验,有

$$4 > 0 \quad \det\begin{bmatrix} 4 & 0 \\ 0 & 10 \end{bmatrix} = 40 > 0 \qquad \det\begin{bmatrix} 4 & 0 & 2 \\ 0 & 10 & 2 \\ 2 & 2 & 2 \end{bmatrix} = 24 > 0$$

故 f''_{X^*} 为正定,在 $\boldsymbol{X}^* = [1, 1, -2]^{\mathrm{T}}$ 处,$f(\boldsymbol{X}^*)$ 为极小。

2.2　有约束条件的函数极值问题

前面讨论函数的极值问题时,向量 $\boldsymbol{X}$ 的各个分量可独立地选择,相互间无约束。本节将讨论 $\boldsymbol{X}$ 的各分量满足一定约束条件的情况。

设具有 n 个变量的多元函数为

$$f(\boldsymbol{X}) = f(x_1, x_2, \cdots, x_n)$$

$\boldsymbol{X}$ 的各分量满足下面的 m 个等式约束方程:

$$g_j(x_1, x_2, \cdots, x_n) = 0 \qquad j = 1, 2, \cdots, m(m < n) \tag{2-13}$$

若能从 m 个约束方程中解出 m 个 $\boldsymbol{X}$ 的分量,即将它们用其他 $n-m$ 个 $\boldsymbol{X}$ 的分量表示,那么 $\boldsymbol{X}$ 中只剩下 $n-m$ 个独立变量,于是问题可化为求 $n-m$ 个变量的多元函数的无约束极值问题。这就是所谓的"消去法"。由于从 m 个方程(一般是非线性方程)求出 m 个分量常常是困难的,故经常采用"拉格朗日乘子法"。为此,对 m 个约束方程,引入 m 个拉格朗日乘子 $\lambda_1, \lambda_2, \cdots, \lambda_m$,并作出一个辅助函数——拉格朗日函数