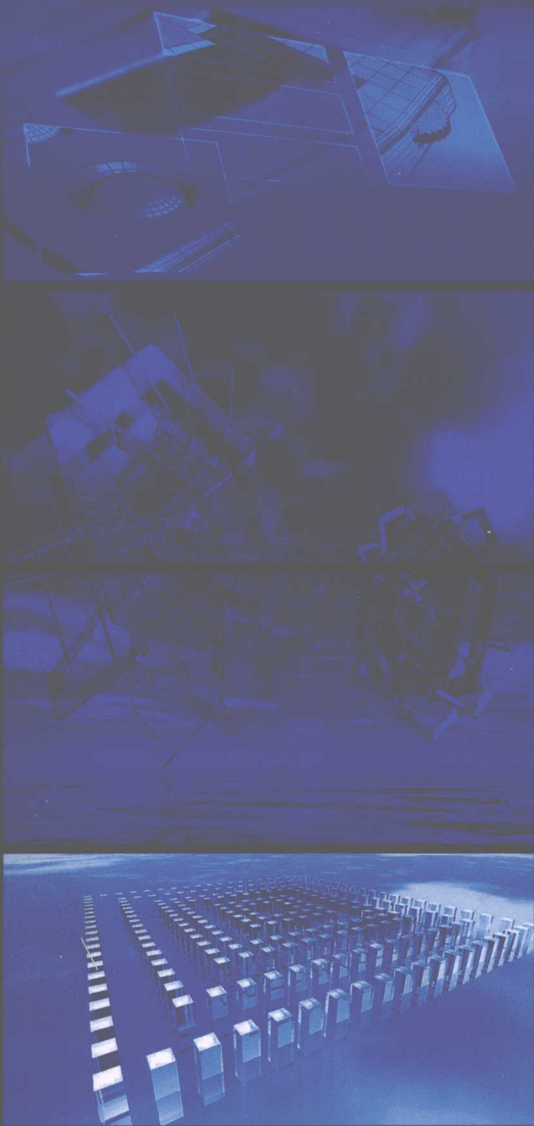
普通高等教育“十一五”国家级规划教材

最优控制理论基础



吕显瑞　黄庆道　编著

科学出版社
www.sciencep.com

普通高等教育“十一五”国家级规划教材

最优控制理论基础

吕显瑞　黄庆道　编著

科学出版社
北京

内容简介

本书共分 7 章，包括最优控制问题、变分法、最大值原理、动态规划、可控性和可观测性、离散控制系统的变分法和最大值原理、线性二次型最优控制问题等.

本书适合数学各专业本科生用作教材，也可供相关专业的老师和学生用作参考书.

图书在版编目(CIP)数据

最优控制理论基础/吕显瑞，黄庆道编著. —北京：科学出版社，2008
普通高等教育"十一五"国家级规划教材
ISBN 978-7-03-021861-2

Ⅰ.最…　Ⅱ.①吕…②黄…　Ⅲ.最佳控制-数学理论-高等学校-教材
Ⅳ.O232

中国版本图书馆 CIP 数据核字(2008)第 063592 号

责任编辑：李鹏奇　王　静　房　阳 / 责任校对：张　琪
责任印制：张克忠 / 封面设计：耕者设计工作室

科学出版社 出版
北京东黄城根北街 16 号
邮政编码：100717
http://www.sciencep.com
双青印刷厂 印刷
科学出版社发行　各地新华书店经销
*
2008 年 7 月第　一　版　开本：B5(720×1000)
2008 年 7 月第一次印刷　印张：10 3/4
印数：1—4 000　字数：201 000

定价：19.00 元

(如有印装质量问题，我社负责调换〈双青〉)

前　言

最优控制理论是现代控制理论中最早发展起来的分支之一. 最优控制理论主要讨论求解最优控制问题的方法和理论,包括最优控制的存在性和唯一性以及最优控制应满足的必要条件等. 最优控制理论始于 20 世纪 50 年代末,其主要标志是前苏联数学家庞特里亚金(L. C. Pontryagin)等提出的"最大值原理". 随着现代科学技术与生产发展的需要以及计算机技术的飞速发展,最优控制理论的应用日益广泛,在工矿企业、交通运输、电力工业、国防工业和国民经济管理等部门有着广泛的应用. 本书系统地介绍了最优控制理论.

全书共分 7 章. 第 1 章通过具体实例介绍了最优控制问题. 第 2 章讨论了变分法的基本概念和变分法在连续系统最优控制中的应用. 第 3 章着重讲述了连续系统最优控制问题的最大值原理及其应用,特别是最小时间控制问题. 第 4 章介绍了动态规划及其在最优控制中的应用. 第 5 章讨论了可控性和可观测性. 第 6 章讲述了离散控制系统最优控制问题的变分法和最大值原理. 第 7 章讨论了线性二次型最优控制问题.

本书是作者在为吉林大学数学学院本科生讲授最优控制理论及应用的讲义基础上,经过多年教学实践不断修改完成的. 吕显瑞编写了第 1 章、第 2 章、4.1 节、4.2 节以及第 7 章,黄庆道编写了第 3 章、4.3 节、第 5 章以及第 6 章. 在编写过程中得到了吉林大学数学学院和数学研究所领导以及李勇教授的热情关怀、帮助和鼓励,在此致以诚挚的谢意! 在编写过程中,参考了许多作者的著作,在此对他们表示衷心的感谢!

由于作者水平有限,书中难免有遗漏和不当之处,恳请广大读者批评指正.

作　者

2008 年 1 月

目　录

第 1 章　最优控制问题 …… 1

1.1　最优控制实例 …… 1

1.2　最优控制问题的数学描述 …… 5

第 2 章　变分法 …… 8

2.1　泛函及其极值 …… 8

2.2　泛函极值的必要条件——欧拉方程 …… 12

2.3　含有多个宗量泛函的极值问题 …… 17

2.4　泛函的条件极值 …… 19

2.5　自由边界条件和横截条件 …… 23

2.6　具终端性能指标的泛函 …… 27

2.7　最优控制问题的变分法 …… 33

2.8　泛函极值曲线的角点条件和充分条件 …… 40

习题 2 …… 43

第 3 章　最大值原理 …… 46

3.1　最大值原理的叙述 …… 46

3.2　最大值原理的证明 …… 51

3.3　最大值原理的应用举例 …… 56

3.4　线性时间最优控制 …… 67

习题 3 …… 79

第 4 章　动态规划 …… 82

4.1　离散型动态规划 …… 82

4.2　动态规划在离散系统最优控制问题中的应用 …… 89

4.3　动态规划在连续系统最优控制问题中的应用 …… 96

习题 4 …… 102

第 5 章　可控性和可观测性 …… 105

5.1　可控性 …… 105

5.2　可观测性 …… 115

5.3　离散系统的可控性和可观测性 …… 118
习题 5 …… 123
第 6 章　离散控制系统的变分法和最大值原理 …… 126
6.1　离散泛函的极值及其变分法 …… 126
6.2　离散控制系统的变分方法 …… 130
6.3　离散控制系统的最大值原理 …… 137
习题 6 …… 139
第 7 章　线性二次型最优控制问题 …… 141
7.1　有限时间的状态调节器问题 …… 141
7.2　无限时间的状态调节器问题 …… 147
7.3　输出调节器问题 …… 151
7.4　跟踪问题 …… 153
7.5　离散系统的线性调节器问题 …… 156
习题 7 …… 162
参考文献 …… 164

第 1 章　最优控制问题

最优控制理论是现代控制理论中最早发展起来的分支之一. 所谓**控制**就是人们用某种方法和手段去影响事件及其运动的进程和轨道，使之朝着有利于控制主体的方向发展. 对于一个给定的受控系统，常常要求找到这样的控制函数，使得在它的作用下，系统从一个状态转移到为设计者希望的另一个状态，且使得系统的某种性能尽可能好. 通常称这种控制问题为**最优控制问题**. **最优控制理论**主要讨论求解最优控制问题的方法和理论，包括最优控制的存在性、唯一性和最优控制应满足的必要条件等. **最优控制理论**始于 20 世纪 50 年代末，其主要标志是前苏联数学家庞特里亚金(L. C. Pontryagin)等提出的“最大值原理”. **最优控制理论**在工矿企业、交通运输、电力工业、国防工业和国民经济管理等部门有着广泛的应用.

本章通过几个具体实例介绍最优控制的基本问题和基本概念以及最优控制问题的数学描述.

1.1　最优控制实例

下面列举几个简单但具有实际应用的例子，它们虽然来自完全不同的领域，但却反映了一个共同的问题——最优控制问题.

例 1.1.1　快速到达问题.

考虑一个机构(如车皮)W，其质量为 m，沿着水平的轨道运动，不考虑空气的阻力和地面对车皮的摩擦力，把车皮看成一个沿着直线运动的质点，$x(t)$表示车皮在 t 时刻的位置，$u(t)$是施加在车皮上的外部控制力，假定车皮的初始位置和速度分别为 $x(0)=x_0$，$\dot{x}(0)=y_0$，希望选择一个控制函数 $u(t)$使车皮在最短时间内到达并静止在坐标原点，即到达坐标原点时速度为零.

根据牛顿第二定律

$$m\ddot{x} = u(t), \quad t > 0, \tag{1.1.1}$$

令 $x=x_1$，$\dot{x}_1=x_2$，则(1.1.1)化为

$$\begin{cases} \dot{x}_1 = x_2, \\ m\dot{x}_2 = u(t), \end{cases} \tag{1.1.2}$$

其中，$x_1(t)$，$x_2(t)$分别表示车皮在 t 时刻的位置和速度，写成向量形式为

$$\begin{cases} \dot{x} = Ax + bu(t), \\ x(0) = (x_0, y_0)^{\mathrm{T}}, \end{cases} \tag{1.1.3}$$

其中，

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 1/m \end{pmatrix}, \quad x(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}.$$

由于技术上的原因，外部推力不可能要多大有多大，它在数量上是有界的，即

$$| u(t) | \leqslant M, \tag{1.1.4}$$

其中，M 是正常数.

$$J(u(t)) = \int_{t_0}^{t_1} 1\mathrm{d}t. \tag{1.1.5}$$

问题是寻找一个满足条件(1.1.4)的控制函数 $u(t)$，把 W 由初态$(x_0, y_0)^{\mathrm{T}}$转移到终态$(0,0)^{\mathrm{T}}$，且使性能指标(1.1.5)达到最小. 任何能达到上述要求的控制函数都称为最优控制. 电梯的快速升降、轧钢机的快速控制和机械振动的快速消振问题都可以用上述问题阐述.

例 1.1.2　国民收入的增长问题.

国民经济收入主要分为两个方面：扩大再生产的积累资金和满足人民生活需要的消费资金. 问题是如何安排积累和消费资金的比例使国民收入得到最快的增长.

用 $x(t)$表示 t 时刻的国民收入，$y(t)$表示用于积累资金的部分，$u(t)=\dfrac{y(t)}{x(t)}$称为积累率. 目的是寻求最优积累率 $u(t)$，使国民收入 $x(t)$增长最快. 国民收入的增长率 $\dot{x}$ 取决于当时的收入总值 $x(t)$和积累率 $u(t)$，即有

$$\dot{x}(t) = f(t, x, u). \tag{1.1.6}$$

根据 $u(t)$的实际含义，$u(t)$应满足

$$0 \leqslant u(t) \leqslant 1.$$

考虑一段时间 T(5 年或 10 年)，使 $x(t)$从初值 x_0 达到尽可能大的 $x(T)$，即

$$x(0) = x_0, \tag{1.1.7}$$

$$\max x(T). \tag{1.1.8}$$

问题归结为在(1.1.6)、(1.1.7)下求 $u(t)$满足(1.1.8). 这等价于它的对偶问题：在固定端点条件 $x(T)=x_1$ 下使 T 最小. 记

$$x(0) = x_0, \quad x(T) = x_1, \tag{1.1.9}$$

$$J(u) = \int_0^T \mathrm{d}t. \tag{1.1.10}$$

国民收入的增长问题的一般提法是在条件(1.1.6)、(1.1.9)下求 $u(t)$ 使性能指标(1.1.10)达到最小.

例 1.1.3 登月艇的软着陆推力控制问题.

登月艇以最小能耗在月球表面进行软着陆的推力控制问题,可以用下列经过简化的问题表示.将登月艇视为一质点,用 $h(t)$ 代表艇距月球表面的距离,以远离月表面的方向为正方向,$v(t)$ 表示艇的速度,$u(t)$ 表示艇上火箭的推力,g 表示月球上的重力加速度,k 为一给定常数,$m(t)$ 为艇的质量,h_0 和 v_0 分别表示艇在初始时刻的高度和速度,M 表示艇在不装燃料时的质量,F 代表燃料的初始质量.登月艇的运动方程为

$$\begin{cases} \dot{h} = v, \\ \dot{v} = -g + \dfrac{u}{m}, \\ \dot{m} = -ku. \end{cases} \tag{1.1.11}$$

初始条件

$$h(t_0) = h_0, \quad v(t_0) = v_0, \quad m(t_0) = M + F. \tag{1.1.12}$$

终端条件

$$h(t_1) = 0, \quad v(t_1) = 0. \tag{1.1.13}$$

这是安全着陆的要求.

作为控制函数的推力 $u(t)$ 满足约束

$$0 \leqslant u(t) \leqslant a, \tag{1.1.14}$$

其中,a 为艇的火箭所能达到的最大推力.

要求燃料最少就是使

$$m(t_0) - m(t_1) = \int_{t_0}^{t_1} -\dot{m}(t)\,\mathrm{d}t \tag{1.1.15}$$

取最小,其中终止时刻 t_1 是待定的.(1.1.15)等价于泛函

$$J(u) = \int_{t_0}^{t_1} u(t)\,\mathrm{d}t. \tag{1.1.16}$$

登月艇的软着陆推力控制问题可以表述为:求满足(1.1.11)~(1.1.14)的 $u(t)$ 使得性能指标泛函(1.1.16)取最小值.

例 1.1.4 基金的最优管理问题.

某基金会得到一笔 60 万元的基金,现将这笔款存入银行,年利率为 10%,该基金计划用 80 年,80 年后要求只剩 0.5 万元用作处理该基金会的结束事宜.根据基金会的需要,每年至少支取 5 万元至多支取 10 万元作为某种奖金.我们的问题

是制定该基金的最优管理策略,即每年支取多少元才能使基金会在 80 年中从银行取出的总金额最大.

令 $x(t)$表示第 t 年存入银行的总钱数,$u(t)$表示第 t 年支取的钱数,则

$$\dot{x}(t)=rx(t)-u(t),\quad r=0.1,\tag{1.1.17}$$

$$x(0)=60,\tag{1.1.18}$$

$$x(80)=0.5,\tag{1.1.19}$$

根据基金会的需要,每年至少支取 5 万元至多支取 10 万元,因此

$$5\leqslant u(t)\leqslant 10,\tag{1.1.20}$$

基金会在 80 年中从银行取出的总金额为

$$J(u)=\int_0^{80}u(t)\mathrm{d}t.\tag{1.1.21}$$

基金的最优管理问题是求满足(1.1.17)～(1.1.20)的 $u(t)$使(1.1.21)中的 $J(u)$取最大值.

例 1.1.5　化学工程问题.

设 $x_1(t),\cdots,x_n(t)$表示反应器中 n 种物质在 t 时刻的数量(浓度),这 n 种物质在反应器中同时发生化学反应. 设反应速率由微分方程组

$$\frac{\mathrm{d}x_i}{\mathrm{d}t}=g_i(x_1,\cdots,x_n,\theta(t),p(t)),x_i(0)=x_i^0,\quad i=1,\cdots,n\tag{1.1.22}$$

来决定,其中,$\theta(t)$,$p(t)$分别表示 t 时刻反应器内的温度和压力,能够在每个时刻控制温度和压力,而温度和压力满足约束条件

$$\theta_1\leqslant\theta(t)\leqslant\theta_2,\quad p_1\leqslant p(t)\leqslant p_2,\tag{1.1.23}$$

其中,$\theta_1,\theta_2,p_1,p_2$ 都是常数,分别表示温度和压力可以达到的最小值和最大值.

假设反应已持续进行了一段时间 T,T 时刻的数量(浓度)为 $x_1(T),\cdots,x_n(T)$. 每种物质的价格为 $c_i\in\mathbf{R},i=1,\cdots,n$. 因此,最终产品的价值是

$$J(p,\theta)=\sum_{i=1}^{n}c_ix_i(T).\tag{1.1.24}$$

问题是寻求$[0,T]$区间上的满足(1.1.22)和(1.1.23)的分段连续函数 $\theta(t)$,$p(t)$,使得(1.1.24)中的 $J(p,\theta)$取最大值.

例 1.1.6　生产计划问题.

设有若干台同样的机器,每台机器可以做两种工作,如果用于做第一种工作,每台每年可获利 3 万元,机器的损坏率为 2/3;如果用于做第二种工作,每台每年可获利 2.5 万元,机器的损坏率为 1/3. 现考虑 3 年的生产周期,试确定如何安排生产计划才可获得最大利润.

设第 k 年可用机器的台数为 $x(k)$，第 k 年分配做第一种工作的机器台数为 $u(k)$，显然，第 $k+1$ 年可用机器的台数 $x(k+1)$ 满足状态方程

$$x(k+1)=\frac{1}{3}u(k)+\frac{2}{3}[x(k)-u(k)], \tag{1.1.25}$$

$u(k)$满足约束条件

$$0\leqslant u(k)\leqslant x(k), \tag{1.1.26}$$

第 k 年获得的利润为

$$R(k)=3u(k)+2.5[x(k)-u(k)]=2.5x(k)+0.5u(k),$$

3 年一共获得的利润为

$$J(u(k))=\sum_{k=0}^{2}R(k)=\sum_{k=0}^{2}[2.5x(k)+0.5u(k)]. \tag{1.1.27}$$

生产计划问题是寻求满足状态方程(1.1.25)和约束条件(1.1.26)的 $u^*(k)(k=0,1,2)$，使(1.1.27)中的目标 $J(u(k))$达到最大.

1.2 最优控制问题的数学描述

前面所举的例子虽然来自不同的领域，但是在数学上都有共同的表述. 将这些最优控制问题作如下数学描述.

1.2.1 控制系统的状态方程

控制系统的**状态变量**是指对事件及其运动起决定作用的量. 控制系统的**控制变量**是指对事件及其运动起控制作用的量. 控制系统的**状态方程**是指描述系统及其运动的方程，其中包含控制变量和状态变量.

令 $x=(x_1,\cdots,x_n)^{\mathrm{T}}\in\mathbf{R}^n$ 表示控制系统的**状态变量**，$u=(u_1,\cdots,u_m)^{\mathrm{T}}\in\mathbf{R}^m$ 表示控制系统的**控制变量**，$t\in I$ 通常表示时间，$f=(f_1,\cdots,f_n)^{\mathrm{T}}$ 是 $X\times U\times I$ 上有定义的 n 维向量函数，则控制系统的**状态方程**通常用一阶常微分方程组

$$\dot{x}=f(t,x(t),u(t)) \tag{1.2.1}$$

来描述.

当 f 不显含 t 时，称(1.2.1)为**定常系统**(或称为**时不变系统**). 当 f 关于 x 和 u 为线性时，称(1.2.1)为**线性系统**. 这时方程可以写成

$$\dot{x}=A(t)x+B(t)u, \tag{1.2.2}$$

其中，$A(t)$为 n 阶方阵，$B(t)$为 n 行 m 列矩阵. 当 A 和 B 与时间 t 无关时，称(1.2.2)为**线性定常系统**或称为**线性自治系统**.

1.2.2　终止状态的目标集

一般来说，控制系统的初始时刻 t_0 和初始状态 $x(t_0)$ 是给定的. 但对控制系统的终止时刻 t_1 和终端状态 $x(t_1)$ 来说，却因问题不同而有不同的要求，通常要求达到一个确定的目标集

$$A = \{x(t_1):x(t_1) \in \mathbf{R}^n, h_1(x(t_1),t_1) = 0, h_2(x(t_1),t_1) \leqslant 0\}.$$

1.2.3　容许控制函数集

在实际问题中，控制变量通常是某种物理量，需要满足有界性等条件，满足这些条件的控制函数，称为**容许控制函数**，它们全体构成一个集合，称为**容许控制函数集**，记为

$$\Omega = \{u(t):u \in \mathbf{R}^m \text{ 并满足某些条件}\}.$$

本书通常要求控制函数是分段连续的.

1.2.4　性能指标

性能指标是指人们对某个控制过程及其结果作出评价的衡量尺度或标准. 在数学上用泛函表示，主要有下面 3 种形式：

(1) 终端型性能指标也称迈耶(Mayer)型性能指标

$$J(u) = \Phi(x(t_1),t_1).$$

(2) 积分型性能指标还称拉格朗日(Lagrange)型性能指标

$$J(u) = \int_{t_0}^{t_1} f_0(t,x(t),u(t))\mathrm{d}t.$$

(3) 混合型性能指标也叫博尔查(Bolza)型性能指标

$$J(u) = \Phi(x(t_1),t_1) + \int_{t_0}^{t_1} f_0(t,x(t),u(t))\mathrm{d}t.$$

1.2.5　最优控制问题的数学描述

最优控制问题就是寻求一容许控制 $u(t)\in\Omega$，使系统的状态从给定的初值 x_0 在终止时刻 $t_1(>t_0)$ 转移到目标集 A，并使性能指标 $J(u)$ 取最大值(或最小值).

若上述最优控制问题有解 $u^*(t)$，则 $u^*(t)$ 称为**最优控制函数**，相应的轨线 $x^*(t)$ 叫做**最优轨线**，而这时的性能指标叫做**最优性能指标**.

1.2.6　离散系统最优控制问题的数学表述

在一些实际问题中，系统的状态变量和控制变量关于时间变量是离散的，这样

的控制系统称为**离散控制系统**.

令 $x(k)=(x_1(k),\cdots,x_n(k))^{\mathrm{T}}\in\mathbf{R}^n,k=k_0,k_0+1,\cdots,k_1$ 表示系统的状态变量,$u(k)=(u_1(k),\cdots,u_m(k))^{\mathrm{T}}\in\mathbf{R}^m,k=k_0,k_0+1,\cdots,k_1-1$ 表示系统的控制变量,则离散控制系统的状态方程为差分方程

$$x(k+1)=\varphi(k,x(k),u(k)),\quad k=k_0,k_0+1,\cdots,k_1-1,$$

其中,$\varphi=(\varphi_1,\cdots,\varphi_n)^{\mathrm{T}}$.

终止状态的目标集为

$$A=\{x(k_1):N_j(k_1,x(k_1))=0,j=1,2,\cdots,l\}.$$

容许控制集为

$$\Omega=\{u(k):u(k)\in\mathbf{R}^m\text{ 并满足某些条件}\}.$$

性能指标为

$$J(u)=\Phi(x(k_1),k_1)+\sum_{k=k_0}^{k_1-1}f_0(k,x(k),u(k)).$$

离散系统的**最优控制问题**就是寻求一容许控制 $u(k)\in\Omega$,使系统的状态从给定的初值 $x(k_0)=x_0$ 在终止时刻 $k_1(>k_0)$转移到目标集 A,并使性能指标 $J(u)$取最大值(或最小值).

若上述最优控制问题有解 $u^*(k)$,则 $u^*(k)$称为**最优控制函数**,相应的轨线 $x^*(k)$叫做**最优轨线**,而这时的性能指标叫做**最优性能指标**.

第2章　变　分　法

变分法(variational calculus)是研究泛函极值的数学方法,早在17世纪末,几何学、力学等领域相继提出了一些泛函极值问题(最速降线问题、最小旋转曲面问题等),导致了变分法的形成和发展.本章介绍变分法及其在最优控制中的应用.

2.1　泛函及其极值

首先给出泛函的定义.

定义 2.1.1　设 Ω 为一函数的集合,若对于每一个函数 $x(t)\in\Omega$,都有一个实数 J 与之对应,则称 J 是定义在 Ω 上的**泛函**,记做 $J(x(t))$. Ω 称为 J 的**容许函数集合**,$x(t)\in\Omega$ 称为**宗量**.

例 2.1.1　对于 xy 平面上过定点 $A(x_1,y_1)$ 和 $B(x_2,y_2)$ 的每一条光滑曲线 $y(x)$,绕 x 轴旋转得一旋转体,旋转体的侧面积是曲线 $y(x)$ 的泛函

$$J(y(x))=\int_{x_1}^{x_2}2\pi y(x)\sqrt{1+\dot{y}^2(x)}\,\mathrm{d}x,$$

容许函数集合可表示为

$$\Omega=\{y(x):y(x)\in C^1[x_1,x_2],y(x_1)=y_1,y(x_2)=y_2\}.$$

第1章中介绍的三个性能指标:

(1) 终端型性能指标也称迈耶型性能指标

$$J(x)=\Phi(x(t_1),t_1).$$

(2) 积分型性能指标还称拉格朗日型性能指标

$$J(x)=\int_{t_0}^{t_1}f_0(t,x(t),\dot{x}(t))\mathrm{d}t.$$

(3) 混合型性能指标也叫博尔查型性能指标

$$J(x)=\Phi(x(t_1),t_1)+\int_{t_0}^{t_1}f_0(t,x(t),\dot{x}(t))\mathrm{d}t.$$

它们都是泛函,并且它们之间可以相互转化.

引进新的函数 $x_0(t)$,它是如下微分方程初值问题的解:

$$\begin{cases}\dot{x}_0(t)=f_0(t,x(t),\dot{x}(t)),\\x_0(t_0)=0,\end{cases}$$

则拉格朗日型性能指标就化为

$$\Phi(x(t_1),t_1) \equiv x_0(t_1) = \int_{t_0}^{t_1} f_0(t,x(t),\dot{x}(t))\mathrm{d}t,$$

变成迈耶型性能指标.

引进函数

$$f_0(t,x(t),\dot{x}(t)) \equiv \frac{\mathrm{d}}{\mathrm{d}t}\Phi(x(t),t) = \Phi_x(x(t),t)\dot{x}(t) + \Phi_t(x(t),t),$$

有

$$\Phi(x(t_1),t_1) = \Phi(x(t_0),t_0) + \int_{t_0}^{t_1} f_0(t,x(t),\dot{x}(t))\mathrm{d}t,$$

其中 $\Phi(x(t_0),t_0)$是已知常数,可以去掉,这样就将迈耶型性能指标转化为拉格朗日型性能指标. 于是,这三种性能指标的任何一种都具有代表性.

为了讨论泛函的连续性,需要定义宗量之间的距离.

定义 2.1.2 设 $x(t),y(t)\in\Omega$,定义它们之间的距离为

$$d_k(x,y) = \max\{\sup|x(t)-y(t)|,\cdots,\sup|x^{(k)}(t)-y^{(k)}(t)|\}, \quad k=0,1,2,\cdots,$$

其中,

$$x(t) = (x_1(t),\cdots,x_n(t))^{\mathrm{T}}, \quad y(t) = (y_1(t),\cdots,y_n(t))^{\mathrm{T}},$$

$$|x(t)-y(t)| = \left(\sum_{i=1}^{n} |x_i(t)-y_i(t)|^2\right)^{\frac{1}{2}},$$

$$\cdots\cdots$$

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

根据上述距离的定义,$d_k<\varepsilon$ 是两个函数之间接近程度的一种描述,$d_0<\varepsilon$ 只要求两个函数坐标之间接近到某种程度,$d_1<\varepsilon$ 不仅要求两个函数的坐标接近而且还要求两个函数的一阶导数之间也要接近. 所以后一个比前一个对两个函数的接近程度要求高. $d_k<\varepsilon$ 表示两个函数具有 k 阶接近度.

宗量 $x_0(t)\in\Omega$ 的 δ 邻域用 $N_i(x_0(t),\delta)$表示,即

$$N_i(x_0(t),\delta) = \{x(t):d_i(x(t),x_0(t)) < \delta\}.$$

定义 2.1.3 设 J 是定义在 Ω 上的泛函,如果对于任给的 $\varepsilon>0$,都可以找到 $\delta>0$,使得当 $d_k(x(t),x_0(t))<\delta$,$x(t)\in\Omega$ 时,就有

$$|J(x(t))-J(x_0(t))| < \varepsilon,$$

则称泛函 J 在 $x_0(x)$处是 k 阶接近的连续泛函.

定义 2.1.4 设 $L(x(t))$是定义在 Ω 上的泛函,如果对任何常数 α,β 都有

$$L(\alpha x(t)+\beta y(t)) = \alpha L(x(t)) + \beta L(y(t)), \quad \forall\, x(t),y(t)\in\Omega,$$

则称 $L(x(t))$是 Ω 上的**线性泛函**.

宗量的改变会引起泛函的变化,把两个宗量之差记为 $\delta x=x_1(t)-x(t)$,则泛函的增量为

$$\Delta J=J(x_1(t))-J(x(t))=J(x+\delta x)-J(x).$$

如同函数的微分是增量的线性主部一样,泛函的变分是泛函的增量的线性主部,即有

定义 2.1.5 设 J 是定义在Ω 上的泛函,如果存在关于 δx 的线性泛函 $L(x,\delta x)$,使得

$$\Delta J=L(x,\delta x)+r(x,\delta x)\cdot\max|\delta x|,$$

其中,$r(x,\delta x)$是 $\max|\delta x|$的高阶无穷小量,即当 $\max|\delta x|\to 0$ 时,有

$$r(x,\delta x)\to 0,$$

称 $L(x,\delta x)$为泛函 $J(x)$的**变分**(variation),记做 $\delta J=L(x,\delta x)$.

泛函的变分是函数的微分概念的推广.

泛函的变分的一个重要形式是它可以表为对参数 α 的导数,即有

$$\delta J(x(t))=\frac{\partial}{\partial\alpha}J(x(t)+\alpha\delta x(t))\Big|_{\alpha=0},$$

这是因为如果变分存在,则增量

$$\Delta J=L(x,\alpha\delta x)+r(x,\alpha\delta x)\cdot\max|\alpha\delta x|,$$

根据 L 和 r 的性质有

$$L(x(t),\alpha\delta x)=\alpha L(x(t),\delta x),$$

$$\lim_{\alpha\to 0}\frac{r(x,\alpha\delta x)\cdot\max|\alpha\delta x|}{\alpha}=\lim_{\alpha\to 0}\frac{\max|\alpha\delta x|}{\alpha\delta x}r(x,\alpha\delta x)\delta x=0,$$

所以

$$\begin{aligned}\frac{\partial}{\partial\alpha}J(x(t)+\alpha\delta x(t))\Big|_{\alpha=0}&=\lim_{\alpha\to 0}\frac{J(x+\alpha\delta x)-J(x)}{\alpha}\\&=\lim_{\alpha\to 0}\frac{L(x,\alpha\delta x)+r(x,\alpha\delta x)\cdot\max|\alpha\delta x|}{\alpha}\\&=\delta J(x).\end{aligned}$$

定义 2.1.6 设 $J(x(t))$是定义在 Ω 上的泛函,$x_0(t)\in\Omega$,若对任意一个 $x(t)\in\Omega$ 都有

$$J(x(t))\geqslant J(x_0(t)),$$

则称 $J(x_0(t))$为 $J(x(t))$的**绝对极小值**(最小值),或者说 $J(x(t))$在 $x_0(t)\in\Omega$ 取

得**绝对极小值**，$x_0(t)\in\Omega$ 称为 $J(x(t))$ 的**极值曲线**.

若存在正数 δ，使得对任意一个 $x(t)\in\Omega\cap N_0(x_0,\delta)$ 都有

$$J(x(t))\geqslant J(x_0(t)),$$

则称 $J(x_0(t))$ 为 $J(x(t))$ 的**强相对极小值**，或者说 $J(x(t))$ 在 $x_0(t)\in\Omega$ 取得**强相对极小值**，$x_0(t)\in\Omega$ 称为 $J(x(t))$ 的**强相对极小值曲线**.

若存在正数 δ，使得对任意一个 $x(t)\in\Omega\cap N_1(x_0,\delta)$ 都有

$$J(x(t))\geqslant J(x_0(t)),$$

则称 $J(x_0(t))$ 为 $J(x(t))$ 的**弱相对极小值**，或者说 $J(x(t))$ 在 $x_0(t)\in\Omega$ 取得**弱相对极小值**，$x_0(t)\in\Omega$ 称为 $J(x(t))$ 的**弱相对极小值曲线**.

可以类似定义绝对极大值、强相对极大值和弱相对极大值.

定义 2.1.7 设 $J(x(t))$ 是定义在 Ω 上的泛函，$x_0(t)\in\Omega$，若对任意一个 $x(t)\in\Omega$ 都有

$$J(x(t))\leqslant J(x_0(t)),$$

则称 $J(x_0(t))$ 为 $J(x(t))$ 的**绝对极大值**(最大值).

若存在正数 δ，使得对任意一个 $x(t)\in\Omega\cap N_0(x_0,\delta)$ 都有

$$J(x(t))\leqslant J(x_0(t)),$$

则称 $J(x_0(t))$ 为 $J(x(t))$ 的**强相对极大值**.

若存在正数 δ，使得对任意一个 $x(t)\in\Omega\cap N_1(x_0,\delta)$ 都有

$$J(x(t))\leqslant J(x_0(t)),$$

则称 $J(x_0(t))$ 为 $J(x(t))$ 的**弱相对极大值**.

由于具有一阶接近度的两个函数必然具有零阶接近度，而反之不一定成立，所以泛函 $J(x_0(t))$ 为 $J(x(t))$ 的强相对极值，则它必为弱相对极值. 反之不一定成立. 由此可知，对于泛函的弱的相对极值的必要条件，必为其强的相对极值的必要条件. 反过来，泛函的强的相对极值的必要条件，不一定为其弱的相对极值的必要条件. 所以，今后在讨论泛函极值必要条件时，总是设 ε 邻域为一级的.

例 2.1.2 设泛函

$$J(y(x))=\int_0^\pi y^2(1-\dot{y}^2)\,\mathrm{d}x,\quad y(0)=y(\pi)=0.$$

试证明函数 $y\equiv0\,(0\leqslant x\leqslant\pi)$ 是泛函 $J(y(x))$ 的弱相对极小值曲线，但不是泛函 $J(y(x))$ 的强相对极小值曲线.

证明 显然 $J(0)=0$. 设 ε 为任意一个小于 1 的正数，N_1 为 $y\equiv0\,(0\leqslant x\leqslant\pi)$ 的一级 ε 邻域，$y(x)$ 为 N_1 中不同于 $y\equiv0\,(0\leqslant x\leqslant\pi)$ 的曲线，则 $y\neq0\,(0\leqslant x\leqslant\pi)$ 时，$y^2(1-\dot{y}^2)>0$，从而