非 线 性 规 划

—理论与算法—

上 册

〔美〕 M·S·巴扎拉 C·M·谢蒂 著

王化存 张春柏 译

贵州大学 印

一九八三年七月

序言

非线性规划是论述一个目标函数在等式约束和不等式约束之下的最优化问题．如果所有这些函数都是线性的，显然，我们就称之为线性规划．否则，就称之为非线性规划．由于线性规划中单纯形法的进一步完善以及高速计算机的出现使得线性规划成为求解不同领域中问题的一个重要工具．然而，许多实际问题由于目标函数和（或）约束函数非线性而不能妥当地表示为一个线性规划问题．因此，人们为了有效地求解非线性问题付出了艰苦的劳动，近二十年来，取得了迅速的进展．本书以逻辑和半封闭形式阐述这些发展．

本书分成三个主要部分，分别论述凸分析，最优性条件，对偶性以及计算方法．凸分析包括凸集和凸函数，它是学习最优化理论的重要内容．最优化的根本目标是研究求解所探讨的问题的有效的计算方法．最优性条件和对偶性不仅可用于研究判定准则，而且也可用以启发计算方法本身．

在本书的准备过程中，特别尽量做到半封闭性．使它既适于作为课本又适于作为参考书．每章都列举有数值例子和图示说明，以帮助读者理解所讨论的概念和方法．此外，每章还附有许多练习．即(1)简单数值问题——巩固课文中所讨论的内容．(2)介绍新材料问题——与课文内容有关．(3)理论练习——供优秀学生使用．每章末，提供了与课文有关的知识介绍，课外参考书和参考资料．这些注释，对读者进一步研究是十分有益的．本书还包含有一个广泛的文献目录．

第一章列举了一些不同工程领域中所出现的问题的例子，这些例子可以看作是非线性规划．包括离散与连续最优控制问题．列举了生产与储存控制和公路设计的实例说明和讨论．另外，还列举有两杆桁架设计和轴承设计的例子．从得到一个二次规划的最优解的观点讨论了电网的平衡条件．提出了

水资源管理中出现的一大型非线性模型．最后，讨论了出现在随机规划和选址理论中的非线性模型．

其余几章归为三个部分．第一部分由第二章和第三章组成．论述凸集和凸函数．第二章讨论凸集的拓扑性质，凸集的分离和支撑，多面体集．多面体集的极点和极方向以及线性规划．．第三章讨论凸函数的性质，包括次微分，凸集上的极大与极小．同时，还讨论了广义凸函数及其相互关系．因为适合于凸函数的非线性规划的算法也可以用于更一般的包括拟凸函数和伪凸函数的函数类

第二部分，包括第四．五．六章，论述最优性条件和对偶性．第四章讨论经典的 Fritz John 和 Kuhn—Tucker 最优性条件，其中，包括等式约束和不等式约束两问题．第五章介绍约束品性．第六章论述 Lagrangian 对偶性和鞍点最优性条件，讨论了对偶理论，对偶函数的性质和求解对偶问题的方法．除了 Lagrangian 对偶性以外，在非线性规划中，还有一些其它的对偶公式，比如，共轭对偶性，极小——极大对偶性，代理对偶和对称对偶性等公式．在这些公式中，Lagrangian 对偶性在算法的发展方面看来是最有前途的．而且，通过每一对偶公式计算其结果是可比较的，鉴于篇幅所限，在课文中，我们选择讨论 Lagrangian 对偶性，而别的对偶公式仅仅放在练习中介绍．

第三部分由第 7 至 11 章组成．不仅讨论求解无约束非线性规划问题的算法，而且也讨论求解有约束非线性规划问题的算法．第七章专门论述视为点到集映射的算法的收敛性定理．这些定理用于证明本书中近年来所提出的算法的收敛性．还简短地讨论了评价算法的准则．第八章论述无约束函数的最优化．我们特别讨论确定线性搜索的几种方法．也讨论多变量函数的极小化问题．不仅讨论使用导数的方法，而且也讨论不使用导数的方法．还讨论了基于共轭概念的方法．第八到 11 章证明所提出的各种算法的收敛性．第 7 章简短地介绍收敛阶的概念，鉴于篇幅和课文水平所限而不做进一步追求．第 9 章讨论罚函数和障碍函数法．在这里，这些方法实际上是把原问题转化为一系列的无约束问题．第 10 章讨论可行方向法，首先假定有一个可行点，

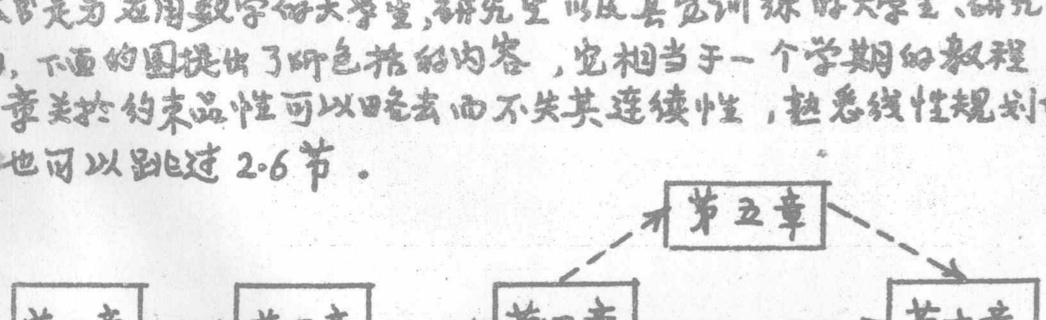
寻找一个可行的改进方向，然后，沿着这个方向，通过极小化目标函数，确定出一个改进的新的可行点。最初的方法是由Zoutendijk所提出，后来为了保证收敛性，由Topkis－Veinott作了修改。Rosen梯度投影法，Wolfe简化梯度法以及Zangwill凸单纯形法都是属于可行方向法这一概念的重要方法，这些方法在第10章一并进行讨论。第11章论述某些特殊的线性约束问题，这些问题可以用稍加修改的线性规划的单纯形法求解，特别，我们讨论二次规划，可分规划和线性分式规划。二次规划是利用本章早先介绍的Lemke互补主元标法来求解。

本书既可以用作非线性规划专题的参考文献，也可以作为运筹学、管理科学、工业工程、应用数学以及涉及分析最优化技术的工程规范方面的课本，所讨论的内容要求数学上熟悉和掌握线性代数和微积分学的知识。为方便读者，附录A归纳了本书时常使用的一些数学论题。

作为课本，本书可用作如下详细叙述的最优化基础的教程以及计算方法的教程，本书也可以用作包括这些内容的连续教程，

1、最优化基础

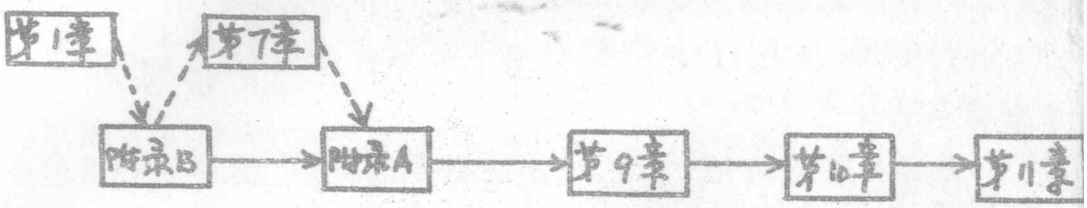
本书是为应用数学的大学生，研究生以及其它训练的大学生、研究生写的，下面的图提供了所包括的内容，它相当于一个学期的教程，第5章关于约束品性可以略去而不失其连续性，熟悉线性规划的读者也可以跳过2.6节。



2、非线性规划中的计算方法

本书是为对于求解非线性规划的算法感兴趣的大学生，研

究生而写的，下面的图提出了所包括的内容．它相当于一个学期的教程，对收敛性不感兴趣的读者可以跳过第7章和第8到第11章中关于收敛性的讨论，对于学习第8到11章 中关于凸分析和最优性条件所必须的最起码的基础知识，为了方便读者把它归纳在附录B中．第1章，给出了非线性规划问题的许多例子，对本教程作了很好的引进，但是，跳过这一章也不失其连续性



（以下略）

Atlanta, Georgia
Janaory. 1979

Mokhtar Bazaraa
C. M. Shetty

3/1

肖荣圭赠

14

目录

第一章

引言

工程师和运筹学工作者们往往都需要求解这样的问题．这些问题的解答可以包括最优设计、稀有资源的最优分配方案或者火箭运行的最优轨道等．过去，关于这些问题的解决方案是一个很大的范围，而认为在这一范围内的方案都是可以采用的．因此，在工程设计中，例如，普遍都会有一个大的安全因子．然而，由于不断的竞争，仅仅取用一可采用的设计方案已不再是妥当的了．另一方面，比如，空间車辆设计，这种可采用的设计本身可能受到限制．因此，实际上需要回答这样一些问题：我们能够做到最有效地利用我们稀有的资源吗？我们能够获得一个最经济的设计吗？在可接受的限制之内我们在进行冒险吗？三十年来，在这些问题的压力之下，由于人们不断的努力探讨，使得在最优化模型和最优化技术方面，取得了很迅速的发展．很幸运，随着大型快速计算机的出现，使得已经提出的最优化技术在使用方面有了坚实可靠的基础．

另一方面，自第二次世界大战以来，对于所求解的问题在其模型和复杂性方面都迅速地增加．因此，有系统的研究成了这一刺激的结果．于是，工程师和运筹学工作者们就被唤醒起来研究同一问题的各个方面，以及研究这些方面的错综复杂的相互关系．其中，一些相互关系甚至不能被充分理解．在一个系统作为一个整体来观察之前，必须理解该系统各部分之间的相互制约关系．由于与统计分析方法相结合的测量技术的发展，对检查我们关于该系统各部分之间相互制约性关系的假设是一个十分重要的方法．

运筹学所研究的范围可以归属或至少部分归属于工业、商业、军事和政府活动中．为了扩大运筹学的研究途径和研究方法以帮助决策者作出决策，早在世界大战以后，运筹学在工业中的应用主要是在线性

规划和统计分析两个方面．这个时候，有效的计标程序和计标机代码对处理这些问题是很有效的．本书论述非线性规划，包括刻划最优解的特征和提出计标程序．

这一章我们引进非线性规划问题，並就这一问题的某些简单情形进行了讨论．我们的目的仅仅是提出非线性规划问题的某些背景，而不打标作深入的研究．

1.1．问题陈述和基本定义

考虑下面的非线性规划问题：

$$
\begin{array}{lll}
\text{Minimize} & f(x) & \\
\text{s.t}^{(*)} & g_i(x)\leq & i=1,\cdots,m \\
 & h_i(x)=0 & i=1,\cdots,l \\
 & x\in X &
\end{array}
$$

其中，f，$g_1,\cdots,g_m$，$h_1,\cdots,h_l$是在E_n上定义的函数，X是E_n中子集和x是n元向量，其分量为$x_1,\cdots,x_n$．求解上面的问题，应该是寻求满足所有约束同时又使函数$f(x)$取极小的变量$x_1,\cdots,x_n$的值．

函数f通常称为目标函数或者判别函数．约束$g_i(x)\leq 0$，$i=1,\cdots,m$称为不等式约束．约束$h_i(x)=0$，$i=1,\cdots,l$称为等式约束．满足所有约束的向量$x\in X$称为问题的可行解．所有可行解的全体组成可行域．非线性规划问题，则是寻找一个可行解$\bar{x}$使得对每一可行点x有$f(x)\geq f(\bar{x})$．这样的点$\bar{x}$称为该问题的最优解或是简单地称为解．

另行说，一个非线性规划问题可以表述为一个极大问题，而不等式约束可以写成$g_i(x)\geq 0$，$i=1,\cdots,m$．在目标函数是线性和包含集X的所有约束可以表示为线性不等式和（或）线性等式的特殊情形下，上面的问题就称为一个线性规划．

为了说明起见，考虑下面的问题：

$$\text{Minimize} \qquad (x_1-3)^2+(x_2-2)^2$$

(*) S.t.是"subject to"的缩写，汉义为"约束条件"——译者

$$\text{s.t.}\quad x_1^2 - x_2 - 3 \le 0$$
$$x_2 - 1 \le 0$$
$$-x_1 \le 0$$

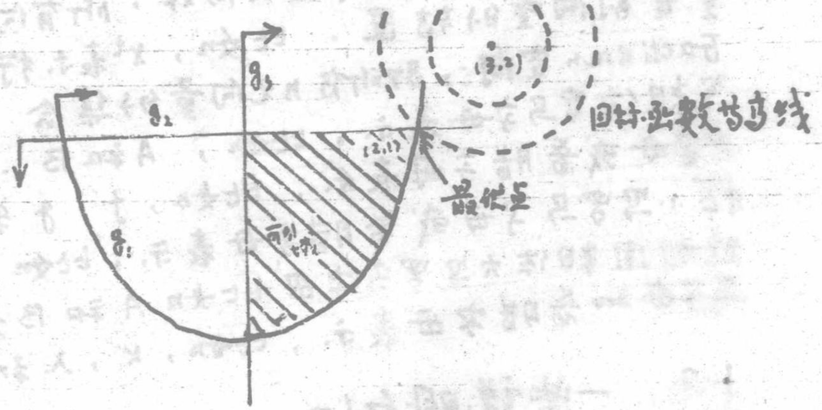


图 1.1　非线性问题的几何解答.

目标函数和三个不等式约束是

$$f(x_1, x_2) = (x_1-3)^2 + (x_2-2)^2$$
$$g_1(x_1, x_2) = x_1^2 - x_2 - 3$$
$$g_2(x_1, x_2) = x_2 - 1$$
$$g_3(x_1, x_2) = -x_1$$

可行区域如图 1.1 所示. 该问题是在可行区域上寻找使得 $(x_1-3)^2+(x_2-2)^2$ 为最小的点. $(x_1-3)^2+(x_2-2)^2=C$ 的点 (x_1, x_2) 是表示以 (3.2) 为中心 $\sqrt{C}$ 为半径的圆, 此圆称为目标函数的等高线. 由于我们希望极小化 C, 于是, 我们应该寻找与可行域相交且具有最小半径的圆. 如图 1.1 所示, 这一最小的圆是 $C=2$ 并且与可行域在点 (2.1) 相交. 所以最优解是 (2.1), 其相应的目标值等于 2.

为了寻找最优解, 上面使用的方法是确定与可行域相交的目标函数为最小的目标等高线. 显然, 这一求解问题的方法, 一般仅仅适用于小问题, 对多于两个变量或有复杂目标函数和约束函数的问题那是不实行的.

记号

本书使用下面的记号，向量用粗体小写罗马字母表示，比如，x，y 和 z。除了特别指出而外，所有向量都是列向量。行向量是列向量的转置。比如，x^t 表示行向量 $(x_1, \cdots, x_n)$。n 维 Euclidian 空间，即所有 n 元向量的集合，用 E_n 表示。矩阵用大写粗体罗马字母表示，比如，A 和 B。纯量值函数用小写罗马字母或希腊字母表示，比如，f，g 和 θ。向量值函数用粗体小写罗马字母或希腊字母表示，比如，g 和 ψ。点到集的映射用粗体大写罗马字母比如 A 和 B 表示。纯量用小写罗马字母和希腊字母表示，比如，k，λ 和 α。

1.2 一些说明例子

这一节我们讨论一些可以建立非线性规划问题的例子，特别讨论如下几方面的最优化问题：

A. 最优控制
B. 结构设计
C. 机械设计
D. 电网络
E. 水资源管理
F. 随机物资分配
G. 设施位置

A 最优控制问题

正如我们即将学习的一个离散控制问题可以表述为一个非线性规划问题，而且一个连续最优控制问题可以用一个非线性规划问题去逼近。因此，本书后面讨论的方法可以用来求解某些最优控制问题。

离散最优控制

考虑一个固定为K个时段的离散最优控制问题．在时段k的开始，该系统用状态向量 y_{k-1} 表示，在时段k末，控制向量 u_k 按照下面的关系把该系统的状态从 y_{k-1} 改变到 y_k．

$$y_k = y_{k-1} + \phi_k(y_{k-1},\ u_k)$$

给出初始状态 y_0，应用控制序列 $u_1,\dots,u_K$ 能够得到一状态向量序列 $y_1,\dots,y_K$，这些状态向量称之为轨道．该过程如图1.2所示．

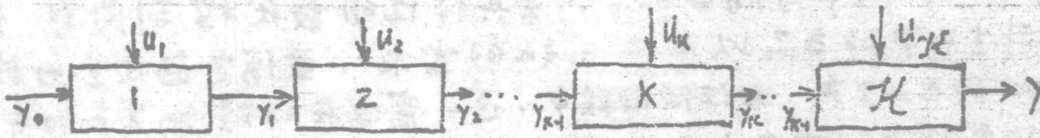


图1.2　离散控制系统说明

控制序列 $u_1,\dots,u_K$ 和状态向量序列 $y_0,\dots,y_K$ 称为是可容许的或可行的，如果它们满足下面的约束：

$$y_k \in Y_k \qquad k=1,\dots,K$$
$$u_k \in U_k \qquad k=1,\dots,K$$
$$\psi(y_0,\dots,y_K,u_1,\dots,u_K) \in D$$

其中 $Y_1,\dots,Y_K$，$U_1,\dots,U_K$ 和 D 都是指定的集合，ψ 是已知的函数，通常称为轨道约束函数．在所有可行控制和所有可行轨道中，通过优化某一目标函数来寻求一最优控制和一个相应的最优轨道．因此，离散控制问题可以陈述如下：

$$\text{Minimize} \quad \alpha(y_0,y_1,\dots,y_K,u_1,\dots,u_K)$$
$$\text{s.t.} \quad y_k = y_{k-1} + \phi_k(y_{k-1},u_k) \qquad k=1,\dots,K$$
$$y_k \in Y_k \qquad k=1,\dots,K$$
$$u_k \in U_k \qquad k=1,\dots,K$$

$$\Psi(y_0,\cdots,y_K,u_1,\cdots,u_K)\in D$$

然把 $y_1,\cdots,y_K,u_1,\cdots,u_K$ 组合成向量 x，并且通过选取适当的 g、h 和 X，则容易验证，上述问题可以表述为 1.1 节引进的非线性规划问题.

生产与储存例子

我们用下面的生产与储存的案例来具体说明最优控制问题. 假设某公司生产某种产品以满足已知的需求，并假定超过总时段 K 时，必须停止生产. 在任何时段中，这一需要在时段初生产的开始可能会碰到储存问题. 而任何时段最大的生产量受到可用设备的生产能力的限制，于是生产量不可能超过 b 单位. 假设当需要产品过剩时，适当的临时工可以雇佣. 然而，不提倡过分的频繁劳动，因而，花费与任何两个连续时段之间劳动力之差的平方成正比，同时，花费也与从一个时段到下一个时段所引起的产品储存量成正比，因此，需要在时段 $1,\cdots,K$ 中寻找劳动力和产品储存量，使得在满足需求之下而总的花费为最小.

在这一问题中有两个状态变量，在时段 K 之末的储存水平 I_K 和劳动力 L_K. 控制变量 u_K 是时段 K 中所获得的劳动力（$u_K<0$ 意味着劳动力按 $-u_K$ 减少.），于是生产与储存问题可以表述如下：

$$
\begin{aligned}
\text{Minimize}\quad & \sum_{K=1}^{K}(c_1u_K^2+c_2I_K) & \\
\text{s.t.}\quad & L_K=L_{K-1}+u_K & K=1,2,\cdots,K\\
& I_K=I_{K-1}+pL_{K-1}-d_K & K=1,\cdots,K\\
& 0\le L_K\le b/p & K=1,\cdots,K\\
& I_K\ge 0 & K=1,\cdots,K
\end{aligned}
$$

其中，初始储存水平 I_0 和初始劳动力 L_0 是已知的. d_K 是在时段 K 中已知的需求. p 是每个工人在任何时段中生产的产

应数．

续最优控制

在离散控制问题的情形下，控制是在离散点上实施．现在我们将考虑在一段固定时间内的连续控制问题．在此，控制函数u定义在一设计的水平[0, T]上．给出初始状态y_0，状态向量y和控制向量u之间的关系由下面的微分方程给出：

$$y(t)=\phi(y(t),u(t)) \qquad t\in[0,T]$$

控制函数和相应的轨迹函数称为是可容许的，如果下面的约束成立：

$$y\in Y$$
$$u\in U$$
$$\psi(y,u)\in D$$

集U的一个典型的例子是定义在[0, T]上，且对t∈[0, T]有$a\le u(t)\le b$的分段连续函数u(t)的集合．最优控制问题可以表述如下；其中，初始状态向量$y(0)=y_0$是给出的．

$$\text{Minimize} \quad \int_0^T \alpha[y(t),u(t)]\,dt$$

$$\text{s.t.} \quad \dot{y}(t)=\phi[y(t),u(t)] \qquad t\in[0,T]$$
$$y\in Y$$
$$u\in U$$
$$\psi(y,u)\in D$$

一个连续最优控制问题可以用一个离散最优控制问题去逼近．特别，假定把所设计的区间[0, T]分成K个时段，使得$K\Delta=T$．用y_k表示$y(K\Delta)$，用u_k表示$u(K\Delta)$，于是上面的问题可以如下逼近，其中初始状态y_0是给出的．

$$\text{Minimize} \quad \sum_{k=1}^{K}\alpha(y_k,u_k)$$

$$
\begin{aligned}
\text{s.t.}\quad & y_K = y_{K-1} + \phi_K(y_{K-1}, u_K) && K=1,\dots,K \\
& y_K \in Y_K && K=1,\dots,K \\
& u_K \in U_K && K=1,\dots,K \\
& \psi(y_0,\dots,y_K, u_1,\dots,u_K) \in D
\end{aligned}
$$

发射火箭的例子

考虑一只火箭在时间 T 内从地面运动到高度 $\bar{y}$ 的问题. 设 $y(t)$ 表示在时刻 t 火箭离地面的高度, 并设 $u(t)$ 表示在时刻 t 火箭在垂直方向上所作出的推力. 假定火箭的质量是 m, 运动方程给定为

$$\ddot{y}(t) + mg = u(t) \qquad t \in [0,T]$$

其中 $\ddot{y}(t)$ 是时刻 t 的加速度, g 是重力加速度. 进而假设在任何时刻火箭所发挥的最大推力不可能超过 b. 如果目标是希望耗费尽可能小的能量使得火箭在时刻 T 到达高度 $\bar{y}$, 于是, 这问题可以表述如下:

$$
\begin{aligned}
\text{Minimize}\quad & \int_0^T |u(t)|\,dt \\
\text{s.t.}\quad & \ddot{y}(t) + mg = u(t) && t\in[0,T] \\
& |u(t)| \le b && t\in[0,T] \\
& y(T) = \bar{y}
\end{aligned}
$$

其中 $y(0)=0$. 由于带二阶微分方程的火箭问题可以变为两个一阶微分方程的等价问题, 只须通过下面的替换做到:

$y_1 = y$, $y_2 = \dot{y}$ 于是 $\ddot{y} + mg = u$ 等价于 $\dot{y}_1 = y_2$ 和 $\dot{y}_2 + mg = u$. 因此上述问题可以重述如下:

$$\text{Minimize}\quad \int_0^T |u(t)|\,dt$$

s.t.

$$\dot{y}_1(t) = y_2(t) \qquad t \in [0,T]$$
$$\dot{y}(t) = u(t) - mg \qquad t \in [0,T]$$
$$|u(t)| \le b \qquad t \in [0,T]$$
$$y(T) = \bar{y}$$

其中 $y_1(0)=y_2(0)=0$. 假定我们把区间 $[0,T]$ 分成 K 个时段. 为简单起见,假定每一时段长为 1. 在时段 k 末,推力、高度和速度分别用 u_k; $y_{1,k}$ 和 $y_{2,k}$ 表示,于是,上面的问题可以用下面的非线性规划逼近,其中 $y_{1,0}=y_{2,0}=0$.

Minimize $\sum_{k=1}^{K} |u_k|$

s.t.

$$y_{1,k} - y_{1,k-1} = y_{2,k-1} \qquad k=1,\dots,K$$
$$y_{2,k} - y_{2,k-1} = u_k - mg \qquad k=1,\dots,K$$
$$|u_k| \le b \qquad k=1,\dots,K$$
$$y_{1,K} = \bar{y}$$

对这一问题和别的连续最优控制问题感兴趣的读者可以参考 Luenberger [1969] 的著作.

公路建筑的例子

假定要在一条起伏的地形上修筑一条公路. 假设构筑的费用与被清除的和填补的土量成正比. 设 T 是公路之长,并设 $c(t)$ 是在任一给定 $t \in [0,T]$ 的地形的已知高度. 寻求描述公路高度 $y(t)$, $t \in [0,T]$ 的方程.

为了避免公路过分倾斜,最大坡度不应当超过 b_1,即,$|\dot{y}(t)| \le b_1$,此外,为了减少公路的崎岖,要求公路坡度的变化率不应该超过 b_2,即,$|\ddot{y}(t)| \le b_2$,还要应该注意端点条件 $y(0)=a$ 和 $y(T)=b$,于是,该问题可以表述如下:

$$\text{Minimize} \quad \int_0^T |y(t) - c(t)|\, dt$$

$$\begin{aligned} \text{S.t.} \quad & |\dot{y}(t)| \le b_1 \quad & t \in [0,T] \\ & |\ddot{y}(t)| \le b_2 \quad & t \in [0,T] \\ & y(0) = a & \\ & y(T) = b & \end{aligned}$$

注意 控制变量是污物总量或者是被清除的污物总量，即 $u(t) = y(t) - c(t)$.

现在，令 $y_1 = y$ 和 $y_2 = \dot{y}$ 并且把公倍分成K段。为方便起见，假定每段之长为1。用 c_K，$y_{1,K}$，$y_{2,K}$ 分别表示 $c(K)$，$y_1(K)$，$y_2(K)$。于是上述问题可以用下面的非线性规划逼近：

$$\text{Minimize} \quad \sum_{k=1}^{K} |y_{1,k} - c_k|$$

$$\begin{aligned} \text{S.t.} \quad & y_{1,k} - y_{1,k-1} = y_{2,k-1} \quad & k = 1, \dots, K \\ & -b_1 \le y_{2,k} \le b_1 \quad & k = 1, \dots, K \\ & -b_2 \le y_{2,k} - y_{2,k-1} \le b_2 \quad & k = 1, \dots, K \\ & y_{1,0} = a & \\ & y_{1,K} = b & \end{aligned}$$

关于这个例子的详细论述，感兴趣的读者可参考 Citron [1969] 的著作。

B 结构设计

结构设计师们传统的工作在于力图探出这样的设计，该设计能够安全地承受所设计的载荷。最优性的概念仅仅隐含在设计的标准规范和设计师的经验之中。最近，提到结构设计，比如，概念结构设计和宇宙航行、过载设计，对最优性的要