Advances in Applied Mathematics 应用数学进展, **2016, 5(4), 773-782**
Published Online November 2016 in Hans. http://www.hanspub.org/journal/aam
http://dx.doi.org/10.12677/aam.2016.54089

Hans汉斯

First-Order Optimality Conditions for a Class of Discrete-Time Optimal Control Problems

Xiaohang Yang[1], Yingtao Xu[2*], Ying Zhang[1], Linyue Du[3]

[1]College of Mathematics, Physics and Information Engineering, Zhejiang Normal University, Jinhua Zhejiang
[2]College of Xingzhi, Zhejiang Normal University, Jinhua Zhejiang
[3]Personnel Department, Zhejiang Normal University, Jinhua Zhejiang
Email: *754273740@qq.com

Received: Nov. 10th, 2016; accepted: Nov. 26th, 2016; published: Nov. 30th, 2016

Abstract

In this paper, we give the Lipschitz condition for a class of discrete-time control problems in which the system has unique solution. Further, we present the control variable transformation function by using the control parameterization method. As a result, the optimal control problem we are considering is converted to a nonlinear differentiable programming problem. Then we put forward a class of first-order optimality conditions for this optimal control problem. Finally, two examples are provided to demonstrate the effectiveness of the proposed first-order optimality conditions.

Keywords

Contraction Fixed Point Theorem, Control Parameterization Method, Optimality Condition

一类离散时间最优控制问题的一阶最优性条件

杨小杭[1]，徐应涛[2*]，张　莹[1]，杜林岳[3]

[1]浙江师范大学数理与信息工程学院，浙江　金华
[2]浙江师范大学行知学院，浙江　金华
[3]浙江师范大学人事处，浙江　金华
Email: *754273740@qq.com

*通讯作者。

文章引用: 杨小杭，徐应涛，张莹，杜林岳. 一类离散时间最优控制问题的一阶最优性条件[J]. 应用数学进展，2016, 5(4): 773-782. http://dx.doi.org/10.12677/aam.2016.54089

收稿日期：2016年11月10日；录用日期：2016年11月26日；发布日期：2016年11月30日

摘　要

本文对一类离散时间的控制问题，提出了存在唯一解的Lipschitz条件，并进一步地，引进控制参数化方法定义控制变量转化函数，将最优控制问题等价转化为非线性可微规划问题，得到了此类最优控制问题的一阶最优性条件。最后，给出两个算例用以验证如上提出的一阶最优性条件。

关键词

压缩不动点定理，控制参数化方法，最优性条件

1. 引言

对于离散时间最优控制问题，应用控制变量参数化方法[1] [2]将最优控制问题转化为非线性规划问题进而求解是目前比较有效的途径。很多学者对如何运用优化算法求解转化后的非线性规划问题进行了广泛研究[3] [4] [5] [6]，然而对于如何应用非线性规划的最优性条件来给出最优控制问题的最优性条件的研究相对较少。

对于控制系统方程解的存在唯一性问题[7] [8] [9]，在[10]中针对一阶微分、差分方程组给出了使解存在唯一的 Lipschitz 条件，并证明了解关于参数的连续依赖性。在控制参数化方法的相关研究中，Teo 等在[1] [2]中详细介绍了控制参数化方法并给出了严格的数学证明，并基于非线性优化算法[11]设计了最优控制软件 MISER [12]。在[1] [4]中给出了离散时间最优控制问题的连续可微条件，保证了目标函数以及约束条件关于控制向量的梯度是存在的，并且基于伴随方程法[1]给出了梯度的具体计算公式。此外，Fritz John 定理和 Kuhn-Tucker 定理是最优化理论里两个重要的定理。[13]讨论了使 Fritz John 点满足 Kuhn-Tucker 条件的 MFCQ。针对一类离散时间最优控制问题。首先，本文给出精确的控制变量转化函数，将离散时间最优控制问题等价转化为非线性可微规划问题。其次，给出具体的 Lipschitz 条件，使该系统在满足这类 Lipschitz 条件下存在唯一解。然后给出更弱的 I-MFCQ，同时给出该问题的一阶最优性条件。最后，给出两个算例应用最优性条件来求解最优控制问题。

2. 离散时间最优控制问题

基于一阶非线性差分方程给出离散时间最优控制问题标准型如下：

$$x\left(k+1\right)=f\left(k,x\left(k\right),u\left(k\right)\right),k\in\mathcal{M};x\left(0\right)=x^{0};\alpha\le u\left(k\right)\le\beta,k\in\mathcal{M},\tag{1}$$

其中，$\mathcal{M}=\{0,1,\cdots,M-1\}$ 为离散时间节点构成的集合，$x=\left[x_1,\cdots,x_n\right]^{\mathrm{T}}\in\mathbb{R}^n$ 为状态变量，$u=\left[u_1,\cdots,u_r\right]^{\mathrm{T}}\in\mathbb{R}^r$ 为控制变量，$x^0\in\mathbb{R}^n$ 为初始状态，$x^0\in\mathbb{R}^n$ 为边界控制量，$f:\mathcal{M}\times\mathbb{R}^n\times\mathbb{R}^r\to\mathbb{R}^n$ 为实值函数，且 $\forall k\in\mathcal{M}$ 有 $f\left(k,\cdot,\cdot\right)$ 关于 x,u 是连续可微的。满足(1)的状态变量称为容许状态，满足(1)的控制变量称为容许控制，并且用 $\mathcal{U}$ 表示容许控制集。

注 1 在满足一定 Lipschitz 条件下，上述系统(1)存在唯一解，具体的 Lipschitz 条件将在下一节中给出. 故对于任意 $k\in\mathcal{M}$，状态变量 $x\left(k\right)$ 可以表示为 $x\left(k\,|\,u\right)$。

记约束条件如下：

$$h_i(u)=\Phi_i\left(x(M\mid u)\right)+\sum_{k=0}^{M-1}\mathcal{L}_i\left(k,x(k\mid u),u(k)\right)=0,\ i=1,2,\cdots,p,\tag{2}$$

$$g_j(u)=\Phi_{p+j}\left(x(M\mid u)\right)+\sum_{k=0}^{M-1}\mathcal{L}_{p+j}\left(k,x(k\mid u),u(k)\right)\le 0,\ j=1,2,\cdots,q,\tag{3}$$

其中，$\Phi_i:\mathbb{R}^n\to\mathbb{R},i=0,\cdots,p+q$ 为连续可微的实值函数；$\mathcal{L}_j:\mathcal{M}\times\mathbb{R}^n\times\mathbb{R}^r\to\mathbb{R}$，$j=0,\cdots,p+q$ 为实值函数，且 $\forall k\in\mathcal{M}$ 有 $\mathcal{L}_j(k,\cdot,\cdot)$ 在 $\mathbb{R}^n\times\mathbb{R}^r$ 上连续可微。

$$J_1(u)=\Phi_0\left(x(M\mid u)\right)+\sum_{k=0}^{M-1}\mathcal{L}_0\left(k,x(k\mid u),u(k)\right),\tag{4}$$

其中，$\Phi_0:\mathbb{R}^n\to\mathbb{R}$ 为连续可微实值函数；$\mathcal{L}_0:\mathcal{M}\times\mathbb{R}^n\times\mathbb{R}^r\to\mathbb{R}$ 为实值函数，并且 $\forall k\in\mathcal{M}$ 有 $\mathcal{L}_0(k,\cdot,\cdot)$ 在 $\mathbb{R}^n\times\mathbb{R}^r$ 上连续可微。称满足(2)-(3)的容许控制为可行控制，用 $\mathcal{F}$ 表示可行控制集。

问题 1 对于离散时间最优控制系统(1)，求解最优控制 $u^*\in\mathcal{F}$ 使之满足约束条件(2)~(3)，并且使性能指标(4)达到最优。

3. 最优控制问题的转换及相关性质

令 $\xi=\left[\xi^0,\xi^1,\cdots,\xi^{M-1}\right]^{\mathrm{T}}\in\mathbb{R}^{M\times r}$，且对 $\forall k\in\mathcal{M}$ 有 $\xi^k\in\mathbb{R}^r,\alpha\le\xi^k\le\beta$。所有 ξ 构成的集合用 $\mathcal{Z}$ 表示。基于控制变量参数化方法[1] [2]构建精确控制变量转化函数如下：

$$\psi^k(\xi)=\sum_{l=0}^{M-1}\xi^l\chi_k(l),\ \forall k\in\mathcal{M},\tag{5}$$

其中 $\psi^k(\xi)$ 为 $\mathcal{Z}\subset\mathbb{R}^M$ 到 $\mathbb{R}$ 上的实值函数，$\chi_k(l)$ 为特征函数，表示如下：

$$\chi_k(l)=\begin{cases}1, & l=k\\ 0, & l\ne k\end{cases}\tag{6}$$

考虑如下系统：

$$y(k+1)=f\left(k,y(k),\psi^k(\xi)\right),k\in\mathcal{M};y(0)=x^0,\tag{7}$$

其中，$\xi\in\mathcal{Z},x^0\in\mathbb{R}^n$ 及 $f:\mathcal{M}\times\mathbb{R}^n\times\mathbb{R}^r\to\mathbb{R}^n$ 为上一节中给出的实值函数。

考虑约束条件如下：

$$\tilde{h}_i(\xi)=\Phi_i\left(y(M\mid\xi)\right)+\sum_{k=0}^{M-1}\mathcal{L}_i\left(k,y(k\mid\xi),\psi^k(\xi)\right)=0,\ i=1,2,\cdots,p,\tag{8}$$

$$\tilde{g}_j(\xi)=\Phi_{p+j}\left(y(M\mid\xi)\right)+\sum_{k=0}^{M-1}\mathcal{L}_{p+j}\left(k,y(k\mid\xi),\psi^k(\xi)\right)\le 0,\ j=1,2,\cdots,q.\tag{9}$$

考虑性能指标如下：

$$J_2(\xi)=\Phi_0\left(y(M\mid\xi)\right)+\sum_{k=0}^{M-1}\mathcal{L}_0\left(k,y(k\mid\xi),\psi^k(\xi)\right),\tag{10}$$

其中，$\Phi_i:\mathbb{R}^n\to\mathbb{R},i=0,1,\cdots,p+q$ 以及 $\mathcal{L}_j:\mathcal{M}\times\mathbb{R}^n\times\mathbb{R}^r\to\mathbb{R},j=0,1,\cdots,p+q$ 为上节定义的实值函数。可知新定义的 $\tilde{h}_i(\xi),\tilde{g}_j(\xi),J_2(\xi)$ 均为 $\mathbb{R}^{M\times r}$ 到 $\mathbb{R}$ 上的实值函数。

问题 2 对于离散时间最优控制系统(7)，求解最优容许控制向量 $\xi^*\in\mathcal{Z}$ 使之满足约束条件(8)~(9)，并且使性能指标(10)达到最优。

定理 1 若 $u(k)=\psi^k(\xi)$，则问题 1 与问题 2 等价。

证明 不妨假设性能指标最小即为最优。先用归纳法证明 $x(\cdot|u)=y(\cdot|\xi)$。由初始条件有 $x(0)=x^0=y(0)$。当 $k=0$ 时，有 $y(1|\xi)=f\left(0,y(0),\psi^0(\xi)\right)=f\left(0,x(0),u(0)\right)=x(1|u)$ 成立。现假设 $k=i-1$ 时成立，即 $y(i|\xi)=x(i|u)$。当 $k=i$ 时，有

$$y(i+1|\xi)=f\left(i,y(i|\xi),\psi^i(\xi)\right)=f\left(i,x(i|u),u(i)\right)=x(i+1|u)$$

成立。综上 $x(\cdot|u)=y(\cdot|\xi)$。

假设 ξ^* 为问题 2 的最优解，取 $u^*(k)=\psi^k\left(\xi^*\right)$。下证 u^* 为问题 1 的最优解。由 u^* 定义，易证 $u^*\in\mathcal{F}$。(反证法)若 u^* 不是问题 1 的最优解，即存在 $\bar{u}\in\mathcal{F}$，使得 $J_1(\bar{u})<J_1\left(u^*\right)$，则有

$$\Phi_0\left(x(M|\bar{u})\right)+\sum_{k=0}^{M-1}\mathcal{L}_0\left(k,x(k|\bar{u}),\bar{u}(k)\right)<\Phi_0\left(x\left(M|u^*\right)\right)+\sum_{k=0}^{M-1}\mathcal{L}_0\left(k,x\left(k|u^*\right),u^*(k)\right),$$

取 $\bar{\xi}^k=\bar{u}(k)$，令 $\bar{\xi}=\left[\bar{\xi}^0,\cdots,\bar{\xi}^{M-1}\right]^{\mathrm{T}}$，有 $\psi^k\left(\bar{\xi}\right)=\bar{\xi}^k=\bar{u}(k)$，且可知 $x(\cdot|\bar{u})=y\left(\cdot|\bar{\xi}\right)$。同理有 $x\left(\cdot|u^*\right)=y\left(\cdot|\xi^*\right)$。故

$$\begin{aligned}J_2\left(\bar{\xi}\right)&=\Phi_0\left(x(M|\bar{u})\right)+\sum_{k=0}^{M-1}\mathcal{L}_0\left(k,x(k|\bar{u}),\bar{u}(k)\right)<\Phi_0\left(x\left(M|u^*\right)\right)+\sum_{k=0}^{M-1}\mathcal{L}_0\left(k,x\left(k|u^*\right),u^*(k)\right)\\&=\Phi_0\left(y\left(M|\xi^*\right)\right)+\sum_{k=0}^{M-1}\mathcal{L}_0\left(k,y\left(k|\xi^*\right),\psi^k\left(\xi^*\right)\right)=J_2\left(\xi^*\right),\end{aligned}$$

有 ξ^* 不是问题 2 的最优解，与假设矛盾。即有 u^* 为问题 1 的最优解。

同理可证若 u^* 为问题 1 的最优解，则存在相应 ξ^* 为问题 2 的最优解。综上问题 1 与问题 2 等价。

下面的定理 2 证明了系统方程解的存在唯一性，命题 1~3 给出了问题 2 的连续可微性。

令 $\bar{f}\left(k,x(k),u(k)\right)=f\left(k,x(k),u(k)\right)-x(k)$，则有

$$x(k+1)=x^0+\sum_{i=0}^{k}\bar{f}\left(i,x(i),u(i)\right),k\in\mathcal{M};\alpha\le u(k)\le\beta,k\in\mathcal{M},\tag{11}$$

易证系统(11)等价于系统(1)，于是若系统(11)存在唯一解 x^*，则 x^* 也为系统(1)唯一解。定义范数 $\|\cdot\|_X$ 为 $\|f\|_X=\sup_{i\in\mathcal{M}}\|f(i)\|_1$，其中 $f:\mathcal{M}\to\mathbb{R}^n$ 且 $\|\cdot\|_1$ 为欧氏距离。易知由如上范数定义的状态函数所构成的空间为完备的 Banach 空间，用 X 表示。

假设 1 $\forall k\in\mathcal{M},u\in\mathcal{U}$ 有

1) 对任意 $x\in X$，存在 $L_1>0$ 使得 $\left\|\bar{f}(k,x,u)\right\|_1\le L_1\left(1+\|x\|_1+\|u\|_1\right)$；

2) 对任意 $x^1,x^2\in X$，存在 $L_2>0$ 使得 $\left\|\bar{f}\left(k,x^1,u\right)-\bar{f}\left(k,x^2,u\right)\right\|_1\le L_2\left\|x^1-x^2\right\|_1$。

定理 2 若系统(11)满足假设 1 条件，并且 $(M+1)L_2<1$，则对任意取定 $u\in\mathcal{U}$ 存在唯一解 $x^*(\cdot|u)$。

证明 记 $\Lambda=\mathbb{R}^n$ 为 n 维向量空间，构造数列 $\{\lambda_n\},\lambda_n\in\Lambda$ 使得 $\left\|\lambda_n-x^0\right\|_1\to 0,(n\to\infty)$。构造函数列 $\{\varphi_n\}$，$\{\psi_n\}$，其中 $\varphi_n\in X,\psi_n\in\mathcal{U}$，使得 $\|\varphi_n-x\|_X\to 0,\|\psi_n-u\|_X\to 0,(n\to\infty)$。定义映射 $\phi:\Lambda\times X\times\mathcal{U}\to X$ 如下

$$\phi(\lambda_n,\varphi_n,\psi_n)(k)=\lambda_n+\sum_{i=0}^{k}\bar{f}\left(i,\varphi_n(i),\psi_n(i)\right),k\in\mathcal{M}.$$

由假设 1)易知 ϕ 是有意义的并且 $x(k+1)=\phi\left(x^0,x,u\right)(k)=x^0+\sum_{i=0}^{k}\bar{f}\left(i,x(i),u(i)\right).$

下证 ϕ 是连续的。由假设 2)及[10]有

$$\left\|\phi\left(\lambda_n,\varphi_n,\psi_n\right)-\phi\left(x^0,x,u\right)\right\|_X \le \left\|\lambda_n-x^0\right\|_1+(M+1)\left(L_2\left\|\varphi_n-x\right\|_X+\left\|\overline{f}\left(\cdot,x,\psi_n\right)-\overline{f}\left(\cdot,x,u\right)\right\|_X\right)。$$

由 f 关于 u 连续可微可知 $\overline{f}$ 也是连续可微的。当 $n\to\infty$，有 $\|\psi_n-u\|_X\to 0$，可知 $\forall i\in\mathcal{M}$ 有 $\left\|\psi_n(i)-u(i)\right\|_1\to 0$，故 $\left\|\overline{f}\left(\cdot,x,\psi_n\right)-\overline{f}\left(\cdot,x,u\right)\right\|_X\to 0$。又由 $\left\|\lambda_n-x^0\right\|_1\to 0,\left\|\varphi_n-x\right\|_X\to 0$，有 $\left\|\phi\left(\lambda_n,\varphi_n,\psi_n\right)-\phi\left(x^0,x,u\right)\right\|_X\to 0$，即 ϕ 连续。

下证 ϕ 是压缩的。对任意 $x^0\in\mathbb{R}^n,x^1,x^2\in X,u\in\mathcal{U}$ 由[10]有

$$\left\|\phi\left(x^0,x^1,u\right)-\phi\left(x^0,x^2,u\right)\right\|_X\le(M+1)L_2\left\|x^1-x^2\right\|_X,$$

由 $(M+1)L_2<1$ 有 ϕ 关于 x 是压缩的。由不动点定理[14]可知 ϕ 存在唯一不动点，即对任意 $u\in\mathcal{U}$ 存在唯一的 $x^*(\cdot|u)\in X$ 使得 $x^*(k+1|u)=\phi\left(x^0,x^*,u\right)(k)=x^0+\sum_{i=0}^{k}\overline{f}\left(i,x^*(i),u(i)\right)$。即(11)存在唯一解 $x^*(\cdot|u)$。

定理 3 若系统(7)满足假设 1 条件，并且 $(M+1)L_2<1$，则对任意取定 $\xi\in\mathcal{Z}$，存在唯一解 $y^*(\cdot|\xi)$。

通过系统方程得出状态变量与控制变量的唯一关系，代入性能指标以及约束条件(8)~(10)中，原最优控制问题即转化为了标准的非线性规划问题。

命题 1 对 $\forall k\in\mathcal{M},u(k)=\psi^k(\xi)$，有 $\psi^k(\xi),y(k|\xi),y(M|\xi)$ 在 $\mathcal{Z}$ 上连续可微。

证明 先证 $\psi^k(\xi)$ 的连续可微性。已知对 $\forall\xi=\left[\xi^0,\xi^1,\cdots,\xi^{M-1}\right]^{\mathrm{T}}\in\mathcal{Z}$ 有

$$\psi^k(\xi)=\sum_{l=0}^{M-1}\xi^l\chi_k(l)=\xi^k,\forall k\in\mathcal{M};\frac{\partial\psi^k}{\partial\xi^i}=\begin{cases}I,i=k,\\\mathbf{0},i\ne k,\end{cases}$$

其中，$I,\mathbf{0}\in\mathbb{R}^r$，可见 ψ^k 在 ξ 处连续可微，又由 ξ 的任意性可知 ψ^k 在 $\mathcal{Z}$ 上连续可微。

取 $y_i=f_i$ 为 f 的第 i 个分量，下证 $\forall i=1,\cdots,n$ 有 $y_i(k|\xi),y_i(M|\xi)$ 在 $\mathcal{Z}$ 上连续可微。

当 $k=0$ 时，由 $y(0|\xi)=x^0$ 为定值，有 $y_i(0|\xi)$ 在 $\mathcal{Z}$ 上连续可微。当 $k=1$ 时，有 $y(1|\xi)=f\left(0,y(0),\psi^0(\xi)\right)$，故

$$\frac{\partial y_i(1|\xi)}{\partial\xi}=\sum_{l=1}^{n}\left(\frac{\partial y_i(1|\xi)}{\partial y_l(0|\xi)}\cdot\frac{\partial y_l(0|\xi)}{\partial\xi}\right)+\frac{\partial y_i(1|\xi)}{\partial\psi^0(\xi)}\cdot\frac{\partial\psi^0(\xi)}{\partial\xi}.$$

由 f 在 $\mathbb{R}^n\times\mathbb{R}^r$ 上连续可微，可知 $y_i=f_i$ 也连续可微，即 $\forall j=1,\cdots,n$，$\frac{\partial y_i(1|\xi)}{\partial y_j(0|\xi)}$ 在 $\mathcal{Z}$ 上存在且连续，再由 $u(k)=\psi^k(\xi)$，有 $\frac{\partial y_i(1|\xi)}{\partial\psi^0(\xi)}$ 在 $\mathcal{Z}$ 上存在且连续；由 $k=0$ 时，有 $y(0|\xi)=x^0$ 在 $\mathcal{Z}$ 上连续可微，即 $\frac{\partial y_j(0|\xi)}{\partial\xi}$ 在 $\mathcal{Z}$ 上连续可微;由 ψ^0 在 $\mathcal{Z}$ 上连续可微，有 $\frac{\partial\psi^0(\xi)}{\partial\xi}$ 在 $\mathcal{Z}$ 上存在且连续。综上，可知 $\frac{\partial y_i(1|\xi)}{\partial\xi}$ 在 $\mathcal{Z}$ 上存在且连续。

现假设 $k=j$ 时成立，于是有 $y_i(j|\xi)$ 在 $\mathcal{Z}$ 上连续可微。同上易证 $\frac{\partial y_i(j+1|\xi)}{\partial\xi}$ 在 $\mathcal{Z}$ 上存在且连续，即 $y_i(j+1|\xi)$ 在 $\mathcal{Z}$ 上连续可微，得证 $\forall k\in\mathcal{M},y(k|\xi)$ 在 $\mathcal{Z}$ 上连续可微。

同理亦有 $y(M|\xi)$ 在 $\mathcal{Z}$ 上连续可微。

由定理 1 及命题 1 易证下面两命题成立。

命题 2 对于问题 2，当取 $u(k)=\psi^k(\xi)$ 时有：

1) 对 $\forall i=0,1,\cdots,p+q$， $\Phi_i\left(y(M\mid\xi)\right)$ 关于 ξ 在 $\mathcal{Z}$ 上连续可微；

2) 对 $\forall j=0,1,\cdots,p+q$， $\forall k\in\mathcal{M}$， $\mathcal{L}_j\left(k,y(k\mid\xi),\psi^k(\xi)\right)$ 关于 ξ 在 $\mathcal{Z}$ 上连续可微。

命题 3 对于问题 2，当取 $u(k)=\psi^k(\xi)$ 时，有 $\tilde{h}_i(\xi),\tilde{g}_j(\xi),J_2(\xi)$ 在 $\mathcal{Z}$ 上连续可微。

注 2 由命题 3 可知，在问题 2 中 $\tilde{h}_i(\xi),\tilde{g}_j(\xi),J_2(\xi)$ 关于 $M\times r$ 维矩阵 ξ 的梯度存在。

下面基于伴随方程法[1]给出具体的梯度计算公式。首先定义哈密顿函数如下：

$$H_i\left(k,y(k),\xi,\lambda^i(k+1)\right)=\mathcal{L}_i\left(k,y(k),\psi^k(\xi)\right)+\left(\lambda^i(k+1)\right)^{\mathrm{T}}f\left(k,y(k),\psi^k(\xi)\right),$$

其中， $\lambda^i(k),i=0,\cdots,p+q$ 由如下伴随方程给出：

$$\left(\lambda^i(k)\right)^{\mathrm{T}}=\frac{\partial H_i\left(k,y(k),\xi,\lambda^i(k+1)\right)}{\partial y(k)},k=M-1,\cdots,1;\left(\lambda^i(M)\right)^{\mathrm{T}}=\left(\frac{\partial\Phi_i\left(y(M)\right)}{\partial y(M)}\right)^{\mathrm{T}}。\tag{12}$$

求解(12)并设解为 $\lambda^i(\cdot\mid\xi)$，可得各函数梯度如下：

$$\frac{\partial J_2(\xi)}{\partial\xi}=\sum_{k=0}^{M-1}\frac{\partial H_0\left(k,y(k),\xi,\lambda^0(k+1)\right)}{\partial\psi^k(\xi)}\frac{\partial\psi^k(\xi)}{\partial\xi},$$

$$\frac{\partial\tilde{h}_i(\xi)}{\partial\xi}=\sum_{k=0}^{M-1}\frac{\partial H_i\left(k,y(k),\xi,\lambda^i(k+1)\right)}{\partial\psi^k(\xi)}\frac{\partial\psi^k(\xi)}{\partial\xi},\quad i=1,2,\cdots,p,$$

$$\frac{\partial\tilde{g}_j(\xi)}{\partial\xi}=\sum_{k=0}^{M-1}\frac{\partial H_{j+p}\left(k,y(k),\xi,\lambda^{j+p}(k+1)\right)}{\partial\psi^k(\xi)}\frac{\partial\psi^k(\xi)}{\partial\xi},\ j=1,2,\cdots,q.$$

4. 一阶最优性条件

对 $\forall i=1,2,\cdots,p,\ j=1,2,\cdots,q,\ A=\left[a_{ij}\right]_{r\times r}\in\mathbb{R}^{r\times r}$ 记：

$$\nabla J_1(u)=\left[\frac{\partial J_1}{\partial u(0)},\frac{\partial J_1}{\partial u(1)},\cdots,\frac{\partial J_1}{\partial u(M-1)}\right]^{\mathrm{T}};\nabla J_2(\xi)=\left[\frac{\partial J_2}{\partial\xi^0},\frac{\partial J_2}{\partial\xi^1},\cdots,\frac{\partial J_2}{\partial\xi^{M-1}}\right]^{\mathrm{T}};$$

$$\nabla h_i(u)=\left[\frac{\partial h_i}{\partial u(0)},\frac{\partial h_i}{\partial u(1)},\cdots,\frac{\partial h_i}{\partial u(M-1)}\right]^{\mathrm{T}};\nabla\tilde{h}_i(\xi)=\left[\frac{\partial\tilde{h}_i}{\partial\xi^0},\frac{\partial\tilde{h}_i}{\partial\xi^1},\cdots,\frac{\partial\tilde{h}_i}{\partial\xi^{M-1}}\right]^{\mathrm{T}};$$

$$\nabla g_j(u)=\left[\frac{\partial g_j}{\partial u(0)},\frac{\partial g_j}{\partial u(1)},\cdots,\frac{\partial g_j}{\partial u(M-1)}\right]^{\mathrm{T}};\nabla\tilde{g}_j(\xi)=\left[\frac{\partial\tilde{g}_j}{\partial\xi^0},\frac{\partial\tilde{g}_j}{\partial\xi^0},\cdots,\frac{\partial\tilde{g}_j}{\partial\xi^{M-1}}\right]^{\mathrm{T}};$$

$$\mathcal{J}(u)=\left\{j:g_j(u)=0,\ j=1,\cdots,q\right\};\ \mathcal{J}(\xi)=\left\{j:\tilde{g}_j(\xi)=0,\ j=1,\cdots,q\right\};$$

$$\tilde{g}(\xi)=\left[\tilde{g}_1(\xi),\cdots,\tilde{g}_p(\xi)\right]^{\mathrm{T}};\ \tilde{h}(\xi)=\left[\tilde{h}_1(\xi),\cdots,\tilde{h}_q(\xi)\right]^{\mathrm{T}};\ D(A)=\left[a_{11},a_{22},\cdots,a_{rr}\right]^{\mathrm{T}}.$$

命题 4 当 $u(k)=\psi^k(\xi)$ 时，有 $\nabla J_1(u)=\nabla J_2(\xi),\nabla h_i(u)=\nabla\tilde{h}_i(\xi),\nabla g_j(u)=\nabla\tilde{g}_j(\xi)$。

证明 对任意 $k\in\mathcal{M}$，由定理 1 有 $x(k\mid u)=y(k\mid\xi)$。故

$$\begin{aligned}\frac{\partial J_1}{\partial u(k)}&=\frac{\partial\Phi_0}{\partial x(M\mid u)}\cdot\frac{\partial x(M\mid u)}{\partial u(k)}+\left(\frac{\partial\mathcal{L}_0}{\partial x(i\mid u)}\cdot\frac{\partial x(i\mid u)}{\partial u(k)}+\frac{\partial\mathcal{L}_0}{\partial u(i)}\cdot\frac{\partial u(i)}{\partial u(k)}\right)\\&=\frac{\partial\Phi_0}{\partial y(M\mid\xi)}\cdot\frac{\partial y(M\mid\xi)}{\partial\xi^k}+\sum_{i=0}^{M-1}\left(\frac{\partial\mathcal{L}_0}{\partial y(i\mid\xi)}\cdot\frac{\partial y(i\mid\xi)}{\partial\xi^k}+\frac{\partial\mathcal{L}_0}{\partial\psi^i(\xi)}\cdot\frac{\partial\psi^i(\xi)}{\partial\xi^k}\right)=\frac{\partial J_2}{\partial\xi^k},\end{aligned}$$

即有 $\nabla J_1(u)=\nabla J_2(\xi)$。同理可证 $\nabla h_i(u)=\nabla\tilde{h}_i(\xi),\nabla g_j(u)=\nabla\tilde{g}_j(\xi)$。

定理 4 若 u^* 为问题 1 的局部最优控制，则存在 $\lambda_0,\lambda_i,i=1,\cdots,p,\mu_j,j\in\mathcal{J}(u^*)$，使得

$$\lambda_0\nabla J_1(u^*)+\sum_{i=1}^{p}\lambda_i\nabla h_i(u^*)+\sum_{j\in\mathcal{J}(u^*)}\mu_j\nabla g_j(u^*)=0, \tag{13}$$

$$\lambda_0\ge 0,\ \ \mu_j\ge 0,\ \ j\in\mathcal{J}(u^*), \tag{14}$$

$$(\lambda_0,\lambda,\mu_{\mathcal{J}})\ne 0, \tag{15}$$

其中，$\lambda=(\lambda_1,\cdots,\lambda_p)^{\mathrm{T}}$，$\mu_{\mathcal{J}}=\{\mu_j,j\in\mathcal{J}(u^*)\}$

证明 已知 u^* 为问题 1 的局部最优控制，取 $\xi^{*,k}=u^*(k)$，令 $\xi^*=\left[\xi^{*,0},\cdots,\xi^{*,M-1}\right]^{\mathrm{T}}$，有 $u^*(k)=\psi^k(\xi^*)$，由定理 1 知 ξ^* 为问题 2 的最优控制.取 $\mathcal{J}(\xi^*)=\left\{j:\tilde{g}_j(\xi^*)=0,j=1,\cdots,q\right\}$，易证当 $u^*(k)=\psi^k(\xi^*)$ 时，$\forall j\in\mathcal{J}(\xi^*)$，有 $j\in\mathcal{J}(u^*)$。由文献[13]中定理 8.11 可知，若 ξ^* 为问题 2 的最优控制，则存在 $\lambda_0,\lambda_i,i=1,\cdots,p,\mu_j$ 以及 $j\in\mathcal{J}(\xi^*)$，使得

$$\lambda_0\nabla J_2(\xi^*)+\sum_{i=1}^{p}\lambda_i\nabla\tilde{h}_i(\xi^*)+\sum_{j\in\mathcal{J}(\xi^*)}\mu_j\nabla\tilde{g}_j(\xi^*)=0;$$

$$\lambda_0\ge 0,\mu_j\ge 0,j\in\mathcal{J}(\xi^*);(\lambda_0,\lambda,\mu_{\mathcal{J}})\ne 0,$$

由命题 4，有 $\nabla J_1(u^*)=\nabla J_2(\xi^*),\nabla h_i(u^*)=\nabla\tilde{h}_i(\xi^*),\nabla g_j(u^*)=\nabla\tilde{g}_j(\xi^*)$。代入上式即有定理 4 成立。

定义 1 [15]在问题 1 中，若 u 满足如下条件：

1) $\nabla g_j(u)^{\mathrm{T}}d<\mathbf{0},j\in\mathcal{J}(u);\nabla h_i(u)^{\mathrm{T}}d=\mathbf{0},i=1,\cdots,p$ 有解 $\tilde{d}\in\mathbb{R}^M$，其中 $\mathbf{0}$ 为 r 维零向量；

2) $\nabla h_i(u),i=1,\cdots,p$ 线性无关。

则称 u 符合 $MFCQ$。

定义 2 在问题 1 中，若 u 满足如下条件：

1) $D\left(\nabla g_j(u)^{\mathrm{T}}d\right)<\mathbf{0},j\in\mathcal{J}(u);D\left(\nabla h_i(u)^{\mathrm{T}}d\right)=\mathbf{0},i=1,\cdots,p$ 有解 $\tilde{d}\in\mathbb{R}^{M\times r}$，其中 $\mathbf{0}$ 为 r 维零向量；

2) $\nabla h_i(u),i=1,\cdots,p$ 线性无关。

则称 u 符合 $I-MFCQ$。

命题 5 若 u 符合 $MFCQ$，则必然符合 $I-MFCQ$。

证明 假设 u 符合 $MFCQ$，则存在 $\tilde{d}\in\mathbb{R}^M$，使 $\mathbf{0}$， $j\in\mathcal{J}(u)$； $\nabla h_i(u)^T\tilde{d}=\mathbf{0},\ i=1,\cdots,p$。当取 $d_l=\tilde{d}$，$l=1,\cdots,r$，令 $d=[d_1,d_2,\cdots,d_r]\in\mathbb{R}^{M\times r}$，于是有 $D\left(\nabla g_j(u)^{\mathrm{T}}d\right)=g_j(u)^{\mathrm{T}}\tilde{d}<\mathbf{0}$，$D\left(\nabla h_i(u)^{\mathrm{T}}d\right)=\nabla h_i(u)^{\mathrm{T}}d=\mathbf{0}$ 成立，即有 u 符合 $MFCQ$。

注 3 由命题 5 可知 $I-MFCQ$ 是基于矩阵给出的，其条件相对于 $MFCQ$ 更弱。

定理 5 若 u^* 为问题 1 的局部最优控制，并且满足 $I-MFCQ$，则存在 $\lambda\in\mathbb{R}^p,\mu\in\mathbb{R}^q$ 使得

$$\nabla J_1(u^*)+\sum_{i=1}^{p}\lambda_i\nabla h_i(u^*)+\sum_{j=1}^{q}\mu_j\nabla g_j(u^*)=0, \tag{16}$$

$$\mu_j g_j(u^*)=0,j=1,2,\cdots,p, \tag{17}$$

$$\mu_j\ge 0,j=1,2,\cdots,q. \tag{18}$$

证明 因u^*是问题1的局部最优控制，故由定理4知(13)~(15)成立。若$\lambda_0=0$，则由(13)有

$$\sum_{i=1}^{p}\lambda_i\nabla h_i\left(u^*\right)+\sum_{j\in\mathcal{J}\left(u^*\right)}\mu_j\nabla g_j\left(u^*\right)=0, \tag{19}$$

其中$\mu_j\ge 0,j\in\mathcal{J}\left(u^*\right)$，且$\mu_j,j\in\mathcal{J}\left(u^*\right),\lambda_i,i=1,\cdots,p$不全为零。以满足$I-MFCQ$的$\tilde{d}$与上式相乘有

$$\sum_{i=1}^{p}\lambda_i\nabla h_i\left(u^*\right)^{\mathrm{T}}\tilde{d}+\sum_{j\in\mathcal{J}\left(u^*\right)}\mu_j\nabla g_j\left(u^*\right)^{\mathrm{T}}\tilde{d}=\mathbf{0},$$

其中，$\mathbf{0}$为$r\times r$维零矩阵。故

$$\sum_{i=1}^{p}\lambda_i D\left(\nabla h_i\left(u^*\right)^{\mathrm{T}}\tilde{d}\right)+\sum_{j\in\mathcal{J}\left(u^*\right)}\mu_j D\left(\nabla g_j\left(u^*\right)^{\mathrm{T}}\tilde{d}\right)=\mathbf{0},$$

其中，$\mathbf{0}$为r维零向量。若存在某$\mu_j>0$，则上式不成立，故必有$\mu_j=0$，$j\in\mathcal{J}\left(u^*\right)$。于是(19)式即为$\sum_{i=1}^{p}\lambda_i\nabla h_i\left(u^*\right)=0$，且$\lambda_i,i=1,\cdots,p$不全为零，这与$\nabla h_i\left(u\right),i=1,\cdots,p$线性无关矛盾。故有$\lambda_0>0$，将$\lambda_0$除遍(13)~(15)，再将$\frac{\lambda_i}{\lambda_0},\frac{\mu_i}{\lambda_0}$视为新的$\lambda_i,\mu_i$，即有定理成立。

在此，称(13)~(15)为Fritz John条件，(16)~(18)为$K-T$条件，并且称满足$K-T$条件的点为$K-T$点。易证下面三个定理成立。

定理6 若u^*为问题1的局部最优控制，且$\nabla h_i\left(u^*\right),i=1,\cdots,p,\nabla g_j\left(u^*\right),j\in\mathcal{J}\left(u^*\right)$线性无关，则$u^*$满足$K-T$条件。

定理7 (充分条件)假设$J_2\left(\xi\right),\tilde{g}\left(\xi\right)$是$\mathbb{R}^M$上的凸函数，$\tilde{h}$是线性函数。若$\xi^*$是问题2的$K-T$点，且$u\left(k\right)=\psi^k\left(\xi\right)$，则$u^*$是问题1的全局最优控制。

定理8 (充分条件)假设$\mathcal{Z}$为凸集，$J_2\left(\xi\right)$为$\mathcal{Z}$上的凸函数。若ξ^*是问题2的$K-T$点，且$u\left(k\right)=\psi^k\left(\xi\right)$，则$u^*$是问题1的全局最优控制。

5. 算例

例1 考虑离散时间控制系统：

$$x\left(k+1\right)=x\left(k\right)+u\left(k\right),k\in\mathcal{M};x\left(0\right)=1;u\left(k\right)\ge 0,k\in\mathcal{M},$$

性能指标为：$J\left(k,x,u\right)=x^2\left(M\right)+\sum_{k=0}^{M-1}\left[-3x\left(k\right)+u^2\left(k\right)-2\chi_1\left(k\right)u^2\left(k\right)\right]$，

约束条件为：$g_l\left(u\right)=\chi_l\left(k\right)u\left(k\right)\ge 0,l=1,2$，

取$\mathcal{M}=\{0,1\},M=2$，求使得性能指标J达到最小的最优控制u。

解 由系统方程有$x\left(0\right)=1,x\left(1\right)=\xi_0+1,x\left(2\right)=\xi_0+\xi_1+1$，故性能指标转化为$J\left(\xi\right)=2\xi_0^2+2\xi_0\xi_1-\xi_0+2\xi_1-5$；约束条件转化为$g_1\left(\xi\right)=-\xi_0\le 0,g_2\left(\xi\right)=-\xi_1\le 0$。

由$\nabla^2 J\left(\xi\right)=\begin{bmatrix}4 & 2\\ 2 & 0\end{bmatrix}$半正定，故$J\left(\xi\right)$是凸函数。由定理7，若存在$\xi^*=\left[\xi_0^*,\xi_1^*\right]^{\mathrm{T}}$使上述问题满足$K-T$条件，并取$u^*\left(k\right)=\xi_k^*$，则有$u^*$为原最优控制问题的全局最优控制。于是，由$K-T$条件可列方程如下：

$$4\xi_0+2\xi_1-1-\mu_1=0,2\xi_0+2-\mu_2=0,\xi_0\mu_1=0,\xi_1\mu_2=0;\xi_0,\xi_1,\mu_1,\mu_2\ge 0.$$

解之得唯一解 $\xi^*=\left[\frac{1}{4},0\right]^{\mathrm{T}},\mu^*=\left[0,\frac{5}{2}\right]^{\mathrm{T}}$。故原最优控制问题有唯一的全局最优控制 $u^*=\left[\frac{1}{4},0\right]^{\mathrm{T}}$。

例 2 考虑离散时间控制系统：

$$x(k+1)=x(k)-u^2(k),k\in\mathcal{M};x(0)=0,$$

性能指标为：$J(k,x,u)=-x(M)+\sum_{k=0}^{M-1}\left\{x(k)-(k+2)\left[u(k)-k-1\right]^2\right\}$，

约束条件为：$g_1(k,x,u)=\sum_{k=0}^{M-1}\left[u(k)-1\right]\le 0;g_2(k,x,u)=\sum_{k=0}^{M-1}\left[(k+1)u(k)-k-1\right]\le 0,$

取 $\mathcal{M}=\{0,1\},M=2$，求使得性能指标 J 达到最小的最优控制 u。

解 取 $\xi_k=u(k)$，对 J 两边取负，即有求解 ξ 使 $J_2(\xi)=-J_1(u)$ 最小。由系统方程有 $x(0)=0$，$x(1)=-\xi_0^2$，$x(2)=-\xi_0^2-\xi_1^2$，故性能指标转化为 $J_2(\xi)=2\xi_0^2-4\xi_0+2\xi_1^2-12\xi_1+14$；约束条件转化为 $g_1(\xi)=\xi_0+\xi_1-2\le 0,g_2(\xi)=\xi_0+2\xi_1-3\le 0$。

$\nabla J_2(\xi)=\left[4\xi_0-4,4\xi_1-12\right]^{\mathrm{T}},\nabla g_1(\xi)=[1,1]^{\mathrm{T}},\nabla g_2(\xi)=[1,2]^{\mathrm{T}}$，并且 $\nabla^2 J_2(\xi)=\begin{bmatrix}4 & 0\\ 4 & 0\end{bmatrix}$ 为半正定矩阵，故 $J_2(\xi)$ 是凸函数。由定理 7，若存在 $\xi^*=\left[\xi_0^*,\xi_1^*\right]^{\mathrm{T}}$ 使上述问题满足 $K-T$ 条件，并取 $u^*(k)=\xi_k^*$，则有 u^* 为最优控制问题的全局最优控制。于是由 $K-T$ 条件可列方程如下：

$$4\xi_0-4+\mu_1+\mu_2=0,4\xi_1-12+\mu_1+2\mu_2=0,(\xi_0+\xi_1-2)\mu_1=0,(\xi_0+2\xi_1-3)\mu_2=0,$$

$$\mu_1,\mu_2\ge 0,(\xi_0+\xi_1-2)\le 0,(\xi_0+2\xi_1-3)\le 0.$$

解之得唯一解 $\xi^*=\left[\frac{1}{5},\frac{7}{5}\right]^{\mathrm{T}},\mu^*=\left[0,\frac{16}{5}\right]^{\mathrm{T}}$。故原最优控制问题有唯一的全局最优控制 $u^*=\left[\frac{1}{5},\frac{7}{5}\right]^{\mathrm{T}}$。

基金项目

浙江省大学生科技创新活动计划(2015R404066)，浙江省教育厅科研项目(Y201534211)。

参考文献 (References)

[1] Teo, K.L., Goh, C.J. and Wong, K.H. (1991) A unified Computational Approach to Optimal Control Problems. Longman Scientific and Technical, Essex.

[2] Lin, Q., Loxton, R. and Teo, K.L. (2014) The Control Parameterization Method for Nonlinear Optimal Control: A Survey. *Journal of Industrial and Management Optimization*, **10**, 275-309. https://doi.org/10.3934/jimo.2014.10.275

[3] Loxton, R., Lin, Q. and Teo, K.L. (2013) Minimizing Control Variation in Nonlinear Optimal Control. *Automatica*, **49**, 2652-2664. https://doi.org/10.1016/j.automatica.2013.05.027

[4] Zhang, Y., Yu, C., Xu, Y. and Teo, K.L. (2016) Minimizing Control Variation in Discrete-Time Optimal Control Problems. *Journal of Computational and Applied Mathematics*, **292**, 292-306. https://doi.org/10.1016/j.cam.2015.07.010

[5] Loxton, R., Teo, K.L. and Rehbock, V. (2008) Optimal Control Problems with Multiple Characteristic Time Points in the Objective and Constraints. *Automatica*, **44**, 2923-2929. https://doi.org/10.1016/j.automatica.2008.04.011

[6] Loxton, R., Teo, K.L. and Rehbock, V. (2011) Robust Suboptimal Control of Nonlinear Systems. *Applied Mathematics and Computation*, **217**, 6566-6576. https://doi.org/10.1016/j.amc.2011.01.039

[7] Elay, S. (2005) An Introduction to Difference Equations. 3 rd Edition, Springer, New York.

[8] 周义仓, 曹慧, 肖燕妮. 差分方程及其应用[M]. 北京: 科学出版社, 2014.

[9] 程金发. 分数阶差分方程理论[M]. 厦门: 厦门大学出版社, 2011.

[10] Peng, F. and Zhan, Z. (2011) Existence Uniqueness and Continuity with Respect to Parameter to Solve First-Order Differential-Difference Dynamic Equations. *Guizhou Science*, **29**, 28-31.

[11] Nocedal, J. and Wright, S.J. (2006) Numerical Optimization. 2nd Edition, Springer, New York.

[12] Jennings, L.S., Fisher, M.E., Teo, K.L., *et al*. (2004) MISER3 Optimal Control Software: Theory and User Manual. University of Western Australia, Perth.

[13] Mangasarian, O.L. (1969) Nonlinear Programming. McGraw-Hill Book Co., New York.

[14] 王声望, 郑维行. 实变函数与泛函分析概要第二册[M]. 第三版. 北京: 高等教育出版社, 2005.

[15] 徐增堃. 数学规划导论[M]. 北京: 科学出版社, 2000.