doi:10. 16055/j. issn. 1672-058X. 2020. 0006. 009

Mayer 型线性最优控制问题的一阶充分条件*

淳 黎, 王 彬

(四川师范大学 数学科学学院,成都 610068)

摘 要:针对目标泛函为 Mayer 型的最优控制问题,在目标函数为伪凸的情形下,证明了当控制系统为线性控制时最优控制的一阶充分条件,同时证明了相应的离散最优控制问题的一阶充分条件;作为应用,通过一阶最优性条件将离散最优控制问题等价地转化为有限维变分不等式问题,并利用伪单调变分不等式的算法给出最优控制的一个数值算例。

关键词:最优控制问题;充分条件;变分不等式;伪凸函数

中图分类号:O232 **文献标志码**:A **文章编号**:1672-058X(2020)06-0056-06

0 引 言

设 $A(\cdot)\in C([0,T];\boldsymbol{R}^{n\times n})$,$B(\cdot)\in C([0,T];\boldsymbol{R}^{n\times m})$;$h:R^n\to\mathbf{R}$ 为给定的一阶连续可微函数,$\boldsymbol{R}^{n\times m}$ 与 $\boldsymbol{R}^{n\times n}$ 分别表示 $n\times m$ 与 $n\times n$ 阶矩阵全体,$C([0,T];\boldsymbol{R}^{n\times n})$,$C([0,T];\boldsymbol{R}^{n\times m})$ 表示 $\boldsymbol{R}^{n\times n}$,$\boldsymbol{R}^{n\times m}$ 矩阵值连续映射全体,记平方可积函数空间 $L^2(0,T;R^m)$ 上的范数与内积分别为 $\|\cdot\|_2$ 与 $[\cdot,\cdot]_{L^2}$,记连续函数空间 $C([0,T];R^n)$ 上的范数为 $\|\cdot\|_\infty$,考虑线性控制系统 :

$$\begin{cases}\dot{x}(t)=A(t)x(t)+B(t)u(t),t\in[0,T]\\ x(0)=x_0\end{cases}\tag{1}$$

Mayer 型性能指标:

$$J(u(\cdot))=h(x(T))$$

其中:$x(T)$ 表示式(1)的解 x 在 T 时的值。设控制集 U 为 R^m 中的有界闭凸集,可行控制集 U_{ad} 定义为

$$U_{ad}:=\{u(\cdot):[0,T]\to U\mid u(\cdot)\text{ Lebesgue 可测}\}$$

本文所考虑的 Mayer 型最优控制问题是求 $\bar{u}(\cdot)\in U_{ad}$,使得

$$J(\bar{u}(\cdot))=\inf_{u(\cdot)\in U_{ad}}J(u(\cdot))\tag{2}$$

满足式(2)的 $\bar{u}(\cdot)$ 称为最优控制,式(1)关于 $\bar{u}(\cdot)$ 的解 $\bar{x}(\cdot)$ 称为最优状态,$(\bar{x}(\cdot),\bar{u}(\cdot))$ 称为最优对。

上述最优控制问题广泛存在于实际工业生产中。求解这类最优控制问题的主要方法之一是求解最优控制的一阶必要条件[1]。然而,类似于微积分中求函数极值,最优控制的一阶必要条件并不一定是充分条件。为使得满足一阶必要条件的允许控制为最优控制,必须对最优控制问题施加适当的凸性条件。现有研究中,一阶充分条件中对最优控制系统的凸性要求比较苛刻,比如,对上述 Mayer 型问题通常要求函数 h 为凸函数[2-3]。过高的凸性要求使得相应的一阶充分条件只适用于结构比较特殊的最优控制问题,从而限制了利用一阶最优性条件求解最优控制这一方法的适用范围。本文在仅假设函数 h 为伪凸函数的条件下证明最优控制问题的一阶必要条件的充分性。因此,与已知一阶充分条件相比,文中所得的一阶充分性条件适用于更广泛的最优控制问题。作为应用,本文将原最优控制

收稿日期:2020-01-02;**修回日期**:2020-03-20.

* **基金项目**:国家自然科学基金资助(11701470).

作者简介:淳黎(1995—),女,四川广元人,硕士研究生,从事最优控制、最优化算法研究.

问题等价地转化为一阶必要条件对应的伪单调变分不等式问题,并利用处理伪单调变分不等式的一类双投影算法[4]给出最优控制的一个简单的数值算例。

1　预备知识

定义 1[5]　设 X 为 Hilbert 空间,其上赋予内积 $[\cdot,\cdot]_X$,$h:X\to\mathbf{R}$ 为一阶可微函数,若

$$h(y)>h(x)\Rightarrow[\nabla h(y),x-y]_X<0,\ \forall x,y\in X$$

则称 h 为定义在 X 上的伪凸函数。

定义 2[5]　设 X 为 Hilbert 空间,K 为 X 中的非空闭凸子集,$F:K\to X$ 称为

(i) 单调映射,若对任意 $x,y\in K$,都有

$$[F(x)-F(y),x-y]_X\geqslant 0$$

(ii) 伪单调映射,若对任意 $x,y\in K$,都有

$$[F(y),x-y]_X\geqslant 0\Rightarrow[F(x),x-y]_X\geqslant 0$$

容易证明,凸函数的导函数为单调映射,伪凸函数的导函数为伪单调映射。

定义 3[4]　设 X 为 Hilbert 空间,其上赋予内积 $[\cdot,\cdot]_X$,K 为 X 中的非空闭凸子集,$F:K\to X$ 为给定的映射,有变分不等式问题 VI(F,K):求 $\bar{x}\in K$,使得

$$[F(\bar{x}),x-\bar{x}]_X\geqslant 0,\ \forall x\in K \tag{3}$$

定义 Hamilton 函数:

$$H(t,x,u,P)=[P,\boldsymbol{A}(\boldsymbol{t})x(\cdot)+\boldsymbol{B}(\boldsymbol{t})u(\cdot)] \tag{4}$$

其中:$(t,x,u,P)\in[0,T]\times R^n\times U\times R^n$。

定理 1 [6]　设$(\bar{x}(\cdot),\bar{u}(\cdot))$为式(2)的最优对,则$(\bar{x}(\cdot),\bar{u}(\cdot))$满足如下 Pontryagin 最大值原理(一阶必要条件):

$$\begin{cases}\dot{\bar{x}}(t)=\boldsymbol{A}(\boldsymbol{t})\bar{x}(t)+\boldsymbol{B}(\boldsymbol{t})\bar{u}(t),t\in[0,T]\\ \dot{P}(t)=-\boldsymbol{A}(\boldsymbol{t})^{\mathrm{T}}\boldsymbol{P}(\boldsymbol{t}),t\in[0,T]\\ \bar{x}(0)=x_0,\boldsymbol{P}(\boldsymbol{T})=-\nabla h(\bar{x}(T))\\ \displaystyle\int_0^T<H_u(t,\bar{x}(t),\bar{u}(t),\boldsymbol{P}(\boldsymbol{t})),v(t)-\bar{u}(t)> \\ \qquad \mathrm{d}t\leqslant 0,\forall v(\cdot)\in U_{ad}\end{cases} \tag{5}$$

其中,$P(\cdot)$称为伴随变量,其满足的方程称为伴随方程。

2　主要结果

下面,在函数 h 为伪凸函数的条件下证明一阶必要条件式(5)的充分性。

定理 2　设控制集 U 为非空闭凸集,函数 h 为一阶连续可微伪凸函数,若$(\bar{x}(\cdot),\bar{u}(\cdot))$满足式(5),则$(\bar{x}(\cdot),\bar{u}(\cdot))$是最优控制问题式(2)的最优对。

证明　下面将分两步证明该结论。

第 1 步　首先证明当 h 是 R^n 上的伪凸函数时,最优控制问题式(2)的性能指标 J 是关于控制 u 在 U_{ad}上的伪凸函数。

任取可行控制 $u_1,u_2\in U_{ad}$,设 $J(u_1)<J(u_2)$,分别记 $x_1(\cdot)$与 $x_2(\cdot)$为控制系统式(1)关于可行控制 u_1,u_2 的解。由性能指标的定义,$h(x_1(T))<h(x_2(T))$,因此,根据 h 是 R^n 上的伪凸函数可得:

$$[\nabla h(x_2(T)),x_1(T)-x_2(T)]<0$$

令 $x^\beta(\cdot)$为控制系统式(1)关于控制 $u_2+\beta(u_1-u_2)$的解。由控制系统式(1)的线性性,易知

$$x^\beta(\cdot)-x_2(\cdot)=\beta(x_1(\cdot)-x_2(\cdot))$$

故由 h 的连续可微性及 Lebesgue 控制收敛定理得:

$$\begin{gathered}<\nabla J(u_2),u_1-u_2>_{L^2}=\\ \lim_{\beta\to 0}\frac{J(u_2+\beta(u_1-u_2))-J(u_2)}{\beta}=\\ \lim_{\beta\to 0}\frac{h(x^\beta(T))-h(x_2(T))}{\beta}=\\ \lim_{\beta\to 0}\int_0^1<\nabla h(x_2(T)+\theta(x^\beta(T)-x_2(T))),\\ \frac{1}{\beta}(x^\beta(T)-x_2(T))>\mathrm{d}\theta=\\ <\nabla h(x_2(T)),x_1(T)-x_2(T)><0\end{gathered} \tag{6}$$

因此,由 u_1,u_2 的任意性及伪凸函数的定义可得性能指标 J 是关于 u 的伪凸函数。

第 2 步　证明一阶必要条件式(5)的充分性。设 $\bar{u}\in U_{ad}$满足一阶必要条件式(5)。由于 U 为 R^m 中的闭凸集,容易验证 U_{ad}为 $L^2(0,T;R^m)$中的闭凸集。对任意可行控制 $u(\cdot)\in U_{ad}$及 $\lambda\in(0,1)$,有

$$u^\lambda=\bar{u}+\lambda(u-\bar{u})\in U_{ad}$$

设 $x^{\lambda}(\cdot)$ 为控制系统(1)关于 u^{λ} 的解,并记 $\delta x(\cdot)=x^{\lambda}(\cdot)-\bar{x}(\cdot)$。由控制系统式(1)的线性性质,$y(\cdot):=\dfrac{\delta x(\cdot)-\bar{x}(\cdot)}{\lambda}$ 是方程:

$$\begin{cases}\dot{y}(t)=\boldsymbol{A}(\boldsymbol{t})y(t)+\boldsymbol{B}(\boldsymbol{t})(u(t)-\bar{u}(t))\\ \qquad t\in[0,T],y(0)=0\end{cases}\tag{7}$$

的解。

类似于第一步的讨论,

$$\lim_{\lambda\to 0^+}\frac{J(u^{\lambda}(T))-J(\bar{u}(T))}{\lambda}=[\nabla h(\bar{x}(T)),y(T)]\tag{8}$$

设 $P(\cdot)$ 为一阶必要条件中伴随方程:

$$\begin{cases}\dot{P}(t)=-\boldsymbol{A}(\boldsymbol{t})^{\mathrm{T}}\boldsymbol{P}(\boldsymbol{t}),t\in[0,T]\\ \boldsymbol{P}(\boldsymbol{T})=-\nabla h(\bar{x}(T))\end{cases}\tag{9}$$

的解。由式(7)与式(9)的对偶关系,有

$$\begin{aligned}&<\nabla h(\bar{x}(T)),y(T)>=<-\boldsymbol{P}(\boldsymbol{T}),y(T)>=\\&-\boldsymbol{P}(\boldsymbol{T})^{\mathrm{T}}y(T)+P(0)y(0)=\\&-\int_0^T\boldsymbol{P}(\boldsymbol{t})^{\mathrm{T}}\dot{y}(t)\mathrm{d}t-\int_0^T\dot{P}(t)y(t)\mathrm{d}t=\\&-\int_0^T<\boldsymbol{P}(\boldsymbol{t}),\boldsymbol{A}(\boldsymbol{t})y(t)+\boldsymbol{B}(\boldsymbol{t})(u(t)-\bar{u}(t))>\mathrm{d}t-\\&\int_0^T<y(t),-\boldsymbol{A}(\boldsymbol{t})^{\mathrm{T}}\boldsymbol{P}(\boldsymbol{t})>\mathrm{d}t=\\&\int_0^T<-\boldsymbol{B}(\boldsymbol{t})^{\mathrm{T}}\boldsymbol{P}(\boldsymbol{t}),u(t)-\bar{u}(t)>\mathrm{d}t\end{aligned}\tag{10}$$

结合式(8)与式(10)可知

$$[\nabla J(\bar{u}),u-\bar{u}]_{L^2}=-\int_0^T[\boldsymbol{B}(\boldsymbol{t})^{\mathrm{T}}\boldsymbol{P}(\boldsymbol{t}),u(t)-\bar{u}(t)]\mathrm{d}t$$

由 $\bar{u}$ 满足一阶必要条件式(5),有

$$[\nabla J(\bar{u}),u-\bar{u}]_{L^2}\geqslant 0$$

因此,利用第一步的结果可得 $J(u)\geqslant J(\bar{u})$,再根据 $u(\cdot)\in U_{ad}$ 的任意性,$\bar{u}(\cdot)$ 为式(2)的最优控制。

3 最优控制的数值解

设 $x(\cdot;u)$ 为式(1)中微分方程关于 $u(\cdot)$ 的解,$\boldsymbol{P}(\cdot;\boldsymbol{u})$ 为式(5)中伴随方程关于 $u(\cdot)$ 和 $x(\cdot;u)$ 的解。定义映射 $F:U_{ad}\to L^2(0,T;R^m)$ 如下:

$$F(u)=-\boldsymbol{B}(\cdot)^{\mathrm{T}}\boldsymbol{P}(\cdot;\boldsymbol{u}),\forall u\in U_{ad}\tag{11}$$

于是,$\bar{u}\in U_{ad}$ 满足式(5)等价于 $\bar{u}$ 是 $L^2(0,T;R^m)$ 上的抽象变分不等式 $\mathrm{VI}(F,U_{ad})$ 的解。因此,由定理 5,当函数 h 为伪凸函数时,求解最优控制问题式(2)等价于求解变分不等式 $\mathrm{VI}(F,U_{ad})$。由于变分不等式问题 $\mathrm{VI}(F,U_{ad})$ 定义在无穷维空间 $L^2(0,T;R^m)$,为获得最优控制问题的数值解,需要对 $\mathrm{VI}(F,U_{ad})$ 进行离散,将其转化为有限维空间上的变分不等式问题。

设 $t_0=0,t_1=\Delta t,\cdots,t_i=i\Delta t,\cdots,t_N=T$,其中 N 为离散区间个数,$\Delta t=\dfrac{T}{N}$ 为划分细度。相应地,记

$$\boldsymbol{u}^N:=(\boldsymbol{u}^N(t_0)^{\mathrm{T}},\cdots,\boldsymbol{u}^N(t_{N-1})^{\mathrm{T}})^{\mathrm{T}}$$
$$\boldsymbol{x}^N:=(\boldsymbol{x}^N(t_0)^{\mathrm{T}},\cdots,\boldsymbol{x}^N(t_N)^{\mathrm{T}})^{\mathrm{T}}$$

由控制系统式(1)的 Euler 差分格式得:对 $i=0,\cdots,N-1$,有

$$\begin{cases}\boldsymbol{x}^N(\boldsymbol{t}_{i+1})=\boldsymbol{x}^N(\boldsymbol{t}_i)+(\boldsymbol{A}(\boldsymbol{t}_i)\boldsymbol{x}^N(\boldsymbol{t}_i)+\boldsymbol{B}(\boldsymbol{t}_i)\boldsymbol{u}^N(\boldsymbol{t}_i))\Delta t\\ x(t_0)=x_0\end{cases}\tag{12}$$

若定义

$$\boldsymbol{U}_{ad}^N:=\underbrace{U\times U\times\cdots,\times U}_{N}$$
$$\boldsymbol{J}^N(\boldsymbol{u}^N)=\boldsymbol{h}(\boldsymbol{x}^N(\boldsymbol{t}_N)),\boldsymbol{u}^N\in\boldsymbol{U}_{ad}^N$$

则最优控制问题式(2)对应的离散化问题为求 $\bar{\boldsymbol{u}}^N\in\boldsymbol{U}_{ad}^N$ 及相应的差分方程式(12)的解 $\boldsymbol{x}^N$,使得

$$\boldsymbol{J}^N(\bar{\boldsymbol{u}}^N)=\inf_{\boldsymbol{u}^N\in\boldsymbol{U}_{ad}^N}\boldsymbol{J}^N(\boldsymbol{u}^N)\tag{13}$$

由参考文献[8]中定理 2 可得如下收敛性结果:

定理 3 设控制集 U 为非空有界闭凸集,函数 h 连续,$\tilde{\boldsymbol{u}}^N$ 为离散优化问题式(13)的解。则

$$\lim_{N\to\infty}\boldsymbol{J}^N(\bar{\boldsymbol{u}}^N)=\inf_{u\in U_{ad}}J(u)$$

类似于式(12),记

$$\boldsymbol{P}^N:=(\boldsymbol{P}^N(\boldsymbol{t}_0)^{\mathrm{T}},\cdots,\boldsymbol{P}^N(\boldsymbol{t}_N)^{\mathrm{T}})^{\mathrm{T}}$$

由倒向 Euler 差分格式,定义 P^N 如下:

$$\begin{cases}\boldsymbol{P}^N(\boldsymbol{t}_i)=\boldsymbol{P}^N(\boldsymbol{t}_{i+1})+\boldsymbol{A}(\boldsymbol{t}_{i+1})^{\mathrm{T}}\boldsymbol{P}^N(\boldsymbol{t}_{i+1})\Delta t,i=0,\cdots,N-2\\ \boldsymbol{P}^N(\boldsymbol{t}_{N-1})=\boldsymbol{P}^N(\boldsymbol{t}_N)=-h_x(x(t_N))\end{cases}\tag{14}$$

对给定的 $\boldsymbol{u}^N$,利用式(12)与式(14)可定义映射 $\boldsymbol{F}^N:\boldsymbol{R}^{Nm}\to\boldsymbol{R}^{Nm}$ 如下:

$$\boldsymbol{F}^N(\boldsymbol{u}^N)=-(\boldsymbol{B}(\boldsymbol{t}_0)^{\mathrm{T}}\boldsymbol{P}^N(\boldsymbol{t}_0),\cdots,\boldsymbol{B}(\boldsymbol{t}_{N-1})^{\mathrm{T}}\boldsymbol{P}^N(\boldsymbol{t}_{N-1}))$$

$$\forall \boldsymbol{u}^N \in \boldsymbol{U}_{ad}^N \tag{15}$$

则变分不等式 VI(F, U_{ad})对应的离散型有限维变分不等式为 VI$(\boldsymbol{F}^N, \boldsymbol{U}_{ad}^N)$:求 $\tilde{\boldsymbol{u}}^N \in \boldsymbol{U}_{ad}^N$使得

$$(\boldsymbol{F}^N(\bar{\boldsymbol{u}}^N), \boldsymbol{u}^N - \bar{\boldsymbol{u}}^N) \geqslant 0, \forall \boldsymbol{u}^N \in \boldsymbol{U}_{ad}^N \tag{16}$$

容易证明式(16)是离散优化问题式(13)对应的一阶必要条件。利用定理 2 的证明方法并结合凸分析的知识容易证明如下结果:

定理 4　若控制集 U 为非空闭凸集,函数 h 为一阶连续可微伪凸函数,则 $\boldsymbol{J}^N(\cdot)$是定义在 $\boldsymbol{U}_{ad}^N$上的伪凸函数,并且 $\bar{\boldsymbol{u}}^N$ 是离散优化问题式(13)的解当且仅当 $\tilde{\boldsymbol{u}}^N$ 是有限维变分不等式式(16)的解。

证明　对任意给定的 $\boldsymbol{u}^N \in \boldsymbol{U}_{ad}^N$,由差分方程式(12)可得

$$\boldsymbol{x}^N(\boldsymbol{t}_n) =$$

$$\boldsymbol{x}^N(t_0) + \sum_{i=0}^{n-1}[\boldsymbol{A}(\boldsymbol{t}_i)\boldsymbol{x}^N(\boldsymbol{t}_i)\Delta t + \boldsymbol{B}(\boldsymbol{t}_i)\boldsymbol{u}^N(\boldsymbol{t}_i)\Delta t]$$

其中,$n=0,1,2,\cdots,N$。

下面,还是分两步来证明。

第 1 步　首先证明 $J^N(\cdot)$为伪凸函数。设 $\boldsymbol{u}_1^N$,$\boldsymbol{u}_2^N \in \boldsymbol{U}_{ad}^N$,并且

$$\boldsymbol{J}^N(\boldsymbol{u}_1^N) < \boldsymbol{J}^N(\boldsymbol{u}_2^N)$$

设 $\boldsymbol{x}_1^N, \boldsymbol{x}_2^N$ 分别是差分方程式(12)关于 $\boldsymbol{u}_1^N, \boldsymbol{u}_2^N$ 的解,则由 $\boldsymbol{J}^N(\cdot)$的定义,知

$$\boldsymbol{J}^N(x_1(t_N)) < \boldsymbol{J}^N(x_2(t_N))$$

由 h 为伪凸函数,得

$$[\nabla h(x_2(t_N)), x_1(t_N) - x_2(t_N)] < 0$$

记 $\boldsymbol{x}_{\beta}^N(\cdot)$为差分方程式(12)关于 $\boldsymbol{u}_{\beta}^N := \boldsymbol{u}_2^N + \beta(\boldsymbol{u}_1^N - \boldsymbol{u}_2^N)$的解,则对$\forall\ n=0,1,2,\cdots,N$,有

$$\frac{1}{\beta}(\boldsymbol{x}_{\beta}^N(t_{n+1}) - \boldsymbol{x}_2^N(t_{n+1})) =$$

$$\frac{1}{\beta}\{x(t_0) + \sum_{i=0}^{n}[\boldsymbol{A}(\boldsymbol{t}_i)\boldsymbol{x}_{\beta}^N(t_i)\Delta t +$$

$$\boldsymbol{B}(\boldsymbol{t}_i)(\boldsymbol{u}_2^N(\boldsymbol{t}_i) + \boldsymbol{\beta}(\boldsymbol{u}_1^N(\boldsymbol{t}_i) - \boldsymbol{u}_2^N(\boldsymbol{t}_i))\Delta t)] - x(t_0) -$$

$$\sum_{i=0}^{n}[\boldsymbol{A}(\boldsymbol{t}_i)\boldsymbol{x}_2^N(\boldsymbol{t}_i)\Delta t + \boldsymbol{B}(\boldsymbol{t}_i)\boldsymbol{u}_2^N(t_i)\Delta t]\} =$$

$$\sum_{i=0}^{n}\left[\boldsymbol{A}(\boldsymbol{t}_i)\frac{\boldsymbol{x}_{\beta}^N(\boldsymbol{t}_i) - \boldsymbol{x}_2^N(\boldsymbol{t}_i)}{\beta}\Delta t + \boldsymbol{B}(\boldsymbol{t}_i)(\boldsymbol{u}_1^N(\boldsymbol{t}_i) - \boldsymbol{u}_2^N(\boldsymbol{t}_i))\Delta t\right]$$

另一方面,有

$$\boldsymbol{x}_1^N(t_{n+1}) - \boldsymbol{x}_2^N(t_{n+1}) = \sum_{i=0}^{n}[\boldsymbol{A}(\boldsymbol{t}_i)(\boldsymbol{x}_1^N(\boldsymbol{t}_i) - \boldsymbol{x}_2^N(\boldsymbol{t}_i))\Delta t + B(\boldsymbol{t}_i)(\boldsymbol{u}_1^N(\boldsymbol{t}_i) - \boldsymbol{u}_2^N(\boldsymbol{t}_i))\Delta t]$$

因此,对$\forall\, n=0,1,2,\cdots,N-1$,有

$$\frac{1}{\beta}(\boldsymbol{x}_{\beta}^N(\boldsymbol{t}_{n+1}) - \boldsymbol{x}_2^N(\boldsymbol{t}_{n+1})) - (\boldsymbol{x}_1^N(\boldsymbol{t}_{n+1}) - \boldsymbol{x}_2^N(\boldsymbol{t}_{n+1})) =$$

$$\boldsymbol{A}(\boldsymbol{t}_i)\sum_{i=0}^{n}\left[\frac{\boldsymbol{x}_{\beta}^N(\boldsymbol{t}_i) - \boldsymbol{x}_2^N(\boldsymbol{t}_i)}{\beta} - (\boldsymbol{x}_1^N(\boldsymbol{t}_i) - \boldsymbol{x}_2^N(\boldsymbol{t}_i))\right]\Delta t = 0$$

因此有

$$[\nabla \boldsymbol{J}^N(\boldsymbol{u}_2^N), \boldsymbol{u}_1^N - \boldsymbol{u}_2^N] =$$

$$\lim_{\beta\to 0}\frac{1}{\beta}[\boldsymbol{J}^N(\boldsymbol{u}_2^N + \boldsymbol{\beta}(\boldsymbol{u}_1^N - \boldsymbol{u}_2^N)) - \boldsymbol{J}^N(\boldsymbol{u}_2^N)] =$$

$$\lim_{\beta\to 0}\frac{1}{\beta}[\boldsymbol{h}(\boldsymbol{x}_{\beta}^N(\boldsymbol{t}_N)) - \boldsymbol{h}(\boldsymbol{x}_2^N(\boldsymbol{t}_N))] =$$

$$[\nabla h(\boldsymbol{x}_2^N(\boldsymbol{t}_N)), \boldsymbol{x}_1^N(\boldsymbol{t}_N) - \boldsymbol{x}_2^N(\boldsymbol{t}_N)] < 0$$

根据 $\boldsymbol{u}_1^N$, $\boldsymbol{u}_2^N$ 的任意性及伪凸函数的定义,$\boldsymbol{J}^N(\cdot)$是 $\boldsymbol{U}_{ad}^N$上的伪凸函数。

第 2 步　接下来,证明 $\bar{\boldsymbol{u}}^N$ 是式(13)的解,当且仅当 $\bar{\boldsymbol{u}}^N$ 是 VI$(\boldsymbol{F}^N, \boldsymbol{U}_{ad}^N)$的解。对 $\forall\ \boldsymbol{u}^N \in \boldsymbol{U}_{ad}^N$,设

$$\boldsymbol{u}_{\gamma}^N := \tilde{\boldsymbol{u}}^N + \gamma(\boldsymbol{u}^N - \bar{u}^N)$$

则由于 U 为凸集,$\boldsymbol{u}_{\gamma}^N \in \boldsymbol{U}_{ad}^N$, $\forall \gamma \in (0,1)$,设 $\boldsymbol{x}_{\gamma}^N(\cdot)$为差分方程式(12)关于 $\boldsymbol{u}_{\gamma}^N$ 的解,记 $\boldsymbol{\delta x}_{\gamma}^N(\cdot) := \boldsymbol{x}_{\gamma}^N - \bar{x}$,由上面的计算,

$$\lim_{\gamma\to 0^+}\frac{\boldsymbol{J}^N(\boldsymbol{u}_{\gamma}^N) - \boldsymbol{J}^N(\bar{\boldsymbol{u}}^N)}{\gamma} =$$

$$<\nabla h(\bar{\boldsymbol{x}}^N(\boldsymbol{t}_N)), \boldsymbol{x}^N(\boldsymbol{t}_N) - \bar{\boldsymbol{x}}^N(\boldsymbol{t}_N)>$$

设 $\boldsymbol{P}^N$ 为差分方程式(14)的解,则

$$<\nabla \boldsymbol{h}(\bar{\boldsymbol{x}}^N(\boldsymbol{t}_N)), \boldsymbol{x}^N(\boldsymbol{t}_N) - \bar{\boldsymbol{x}}^N(\boldsymbol{t}_N)> =$$

$$<-\boldsymbol{P}^N(\boldsymbol{t}_N), \sum_{i=0}^{N-1}[\boldsymbol{A}(\boldsymbol{t}_i)(\boldsymbol{x}^N(\boldsymbol{t}_i) - \bar{\boldsymbol{x}}^N(\boldsymbol{t}_i))\Delta t +$$

$$\boldsymbol{B}(\boldsymbol{t}_i)(\boldsymbol{u}^N(\boldsymbol{t}_i) - \tilde{\boldsymbol{u}}^N(\boldsymbol{t}_i))\boldsymbol{\Delta t}]> =$$

$$\sum_{i=0}^{N-1}<-P^N(t_i), \boldsymbol{A}(\boldsymbol{t}_i)(\boldsymbol{x}^N(\boldsymbol{t}_i) - \bar{\boldsymbol{x}}^N(\boldsymbol{t}_i))\boldsymbol{\Delta t} +$$

$$\boldsymbol{B}(\boldsymbol{t}_i)(\boldsymbol{u}^N(\boldsymbol{t}_i) - \bar{u}^N(\boldsymbol{t}_i))\Delta t> +$$

$$\sum_{i=0}^{N-1}<\boldsymbol{P}^N(\boldsymbol{t}_i) - \boldsymbol{P}^N(\boldsymbol{t}_N), \boldsymbol{A}(\boldsymbol{t}_i)(\boldsymbol{x}^N(\boldsymbol{t}_i) - \bar{\boldsymbol{x}}^N(\boldsymbol{t}_i))\boldsymbol{\Delta t} +$$

$$\boldsymbol{B}(\boldsymbol{t}_i)(\boldsymbol{u}^N(\boldsymbol{t}_i) - \bar{\boldsymbol{u}}^N(\boldsymbol{t}_i))\boldsymbol{\Delta t}>$$

由于

$$\boldsymbol{P}^N(\boldsymbol{t}_{N-1}) = \boldsymbol{P}^N(\boldsymbol{t}_N), \boldsymbol{P}^N(\boldsymbol{t}_i) = \boldsymbol{P}^N(\boldsymbol{t}_{N-1}) +$$

$$\sum_{j=i+1}^{N-1}\boldsymbol{A}(\boldsymbol{t}_j)^{\mathrm{T}}\boldsymbol{P}^N(\boldsymbol{t}_j)\Delta t, i = 0,\cdots,n-2$$

因此

$$\sum_{i=0}^{N-1}<\boldsymbol{P}^N(\boldsymbol{t}_i) - \boldsymbol{P}^N(\boldsymbol{t}_N), \boldsymbol{A}(\boldsymbol{t}_i)(\boldsymbol{x}^N(\boldsymbol{t}_i) -$$

$$\bar{\boldsymbol{x}}^N(t_i))\Delta t + \boldsymbol{B}(t_i)(\boldsymbol{u}^N(t_i) - \bar{\boldsymbol{u}}^N(t_i))\Delta t >=$$

$$\sum_{i=0}^{N-1} < \sum_{j=i+1}^{N-1} \boldsymbol{A}(t_j)^{\mathrm{T}}\boldsymbol{P}^N(t_j)\Delta t, \boldsymbol{A}(t_i)(\boldsymbol{x}^N(t_i) -$$

$$\bar{\boldsymbol{x}}^N(t_i))\Delta t + \boldsymbol{B}(t_i)(\boldsymbol{u}^N(t_i) - \bar{\boldsymbol{u}}^N(t_i))\Delta t >=$$

$$\sum_{j=1}^{N-1} < \boldsymbol{A}(t_j)^{\mathrm{T}}\boldsymbol{P}^N(t_j)\Delta t, \sum_{i=0}^{j-1}[\boldsymbol{A}(t_i)(\boldsymbol{x}^N(t_i) -$$

$$\bar{\boldsymbol{x}}^N(t_i))\Delta t + \boldsymbol{B}(t_i)(\boldsymbol{u}^N(t_i) - \bar{\boldsymbol{u}}^N(t_i))\Delta t] >=$$

$$\sum_{j=1}^{N-1} < \boldsymbol{A}(t_j)^{\mathrm{T}}\boldsymbol{p}^N(t_j)\Delta t, \boldsymbol{x}^N(t_j) - \bar{\boldsymbol{x}}^N(t_j) >=$$

$$\sum_{j=1}^{N-1} < \boldsymbol{p}^N(t_j), \boldsymbol{A}(t_j)(\boldsymbol{x}^N(t_j) - \bar{\boldsymbol{x}}(t_j))\Delta t >$$

综上可得

$$< \nabla \boldsymbol{J}^N(\tilde{\boldsymbol{u}}^N), \boldsymbol{u}^N - \bar{\boldsymbol{u}}^N > = \lim_{\gamma\to 0^+}\frac{\boldsymbol{J}^N(\boldsymbol{u}_\gamma^N) - \boldsymbol{J}^N(\bar{\boldsymbol{u}}^N)}{\gamma} =$$

$$< \nabla \boldsymbol{h}(\bar{\boldsymbol{x}}(t_N)), \boldsymbol{x}(t_N) - \bar{\boldsymbol{x}}(t_N) >=$$

$$\sum_{i=0}^{N-1} < -\boldsymbol{P}^N(t_i), \boldsymbol{A}(t_i)(\boldsymbol{x}^N(t_i) - \bar{\boldsymbol{x}}^N(t_i))\Delta t +$$

$$\boldsymbol{B}(t_i)(\boldsymbol{u}^N(t_i) - \bar{\boldsymbol{u}}^N(t_i))\Delta t > +$$

$$\sum_{j=1}^{N-1} < \boldsymbol{P}^N(t_j), \boldsymbol{A}(t_j)(\boldsymbol{x}^N(t_j) - \bar{\boldsymbol{x}}^N(t_j))\Delta t >=$$

$$\sum_{i=0}^{N-1} < -\boldsymbol{B}(t_i)^{\mathrm{T}}\boldsymbol{P}^N(t_i), \boldsymbol{u}^N(t_i) - \bar{\boldsymbol{u}}^N(t_i) > \Delta t =$$

$$< \boldsymbol{F}^N(\bar{\boldsymbol{u}}^N), \boldsymbol{u}^N - \bar{\boldsymbol{u}}^N > \Delta t$$

故若 $\bar{\boldsymbol{u}}^N$ 为式(13)的最优解,则 $[\nabla \boldsymbol{J}^N(\bar{\boldsymbol{u}}^N), \boldsymbol{u}^N - \bar{\boldsymbol{u}}^N] \geqslant 0, \forall \boldsymbol{u}^N \in \boldsymbol{U}_{ad}^N$,它等价于 $\bar{\boldsymbol{u}}^N$ 满足变分不等式 $[\boldsymbol{F}^N(\bar{\boldsymbol{u}}^N), \boldsymbol{u}^N - \bar{\boldsymbol{u}}^N] \geqslant 0, \forall\ \boldsymbol{u}^N \in \boldsymbol{U}_{ad}^N$。

另一方面,若

$$[\boldsymbol{F}^N(\bar{\boldsymbol{u}}^N), \boldsymbol{u}^N - \bar{\boldsymbol{u}}^N] \geqslant 0, \forall\ \boldsymbol{u}^N \in \boldsymbol{U}_{ad}^N$$

则有

$$[\nabla \boldsymbol{J}^N(\bar{\boldsymbol{u}}^N), \boldsymbol{u}^N - \bar{\boldsymbol{u}}^N] \geqslant 0$$

由 $\boldsymbol{J}^N(\cdot)$ 为伪凸函数,有

$$\boldsymbol{J}^N(\boldsymbol{u}^N) \geqslant \boldsymbol{J}^N(\bar{\boldsymbol{u}}^N), \forall\ \boldsymbol{u}^N \in \boldsymbol{U}_{ad}^N$$

故式(16)是 $\bar{\boldsymbol{u}}^N$ 是式(13)的最优解的充要条件。

由定理 7,利用凸分析中的经典结果可证 $\boldsymbol{F}^N(\cdot)$ 为定义在 $\boldsymbol{U}_{ad}^N$ 上的伪单调映射。因此,根据定理 6 与定理 7 的结论,可以利用求解伪单调变分不等式的算法求解最优控的数值近似解。下面给出一个简单的算例。

例 1 设 $T=1, n=m=1$,控制集 $U=[0,1]$。考虑如下控制系统:

$$\begin{cases} \dot{x}(t) = x(t) + 5u(t), t \in (0,1) \\ x(0) = 1 \end{cases}$$

及性能指标

$$J(u(\cdot)) = \ln(x(T))$$

容易证明 $\ln(\cdot)$ 为伪凸函数。由方程的结构及对数函数的单调性易知该最优控制问题的最优解为 $\bar{u}(t) \equiv 0$。

利用例 1 中对应的离散有限维变分不等式及文献[4]中关于伪单调变分不等式的双投影算法可得最优控制的数值解。为此,选取初始点 $\boldsymbol{u}_0^N = (1, \cdots, 1)^{\mathrm{T}}$, $N=100$。在 Windows7 系统,处理器为 nter(R) Core(TM) i5-3337U, CPU 为 4,核为 1.8 GHz 的计算机上使用版本为 R2014a 的 MATLAB 进行数值实验。若设置误差为 10^{-6},当迭代次数 iter = 11 时算法停止,程序运行时间 $t=0.343\ 202$ s。其每一次迭代所得控制的函数图像变化如图 1 所示。

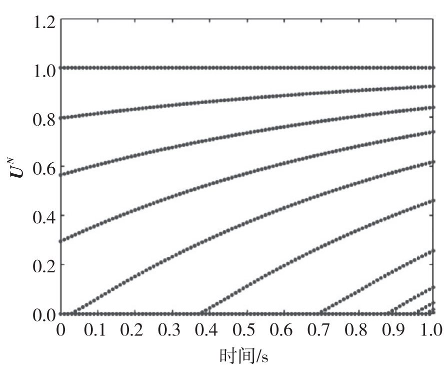

图 1 时间-控制图

Fig. 1 Time-control diagram

致谢:非常感谢四川师范大学数学科学学院张海森老师对本文提供的指导和建议。

参考文献(References):

[1] TEO K L, GOH C J, WONG K H. A Unified Computational Approach to Optimal Control Problems[M]. New York: Longman Scientific and Technical, 1991

[2] 雍炯敏,楼红卫. 最优控制理论简明教程[M]. 北京:高等教育出版社,2006
YONG J M, LOU H W. A Concise Tutorial of Optimal Control Theory[M]. Beijing: Higher Education Press, 2006(in Chinese)

[3] 赵清梅, 张俊容. 最优控制问题的 Tikhonov 正则化[J]. 西南大学学报(自然科学版),2019,41(5):59—63
ZHAO Q M, ZHANG J R. The Tikhonov Regularization of

Optimal Control Problem [J]. Journal of Southwest University(Natural Science Editen), 2019,41(5): 59—63(in Chinese)

[4] HE Y R. A New Double Projection Algorithm for Variational Inequalities[J]. Journal of Computational and Applied Mathematics, 2006, 185(1): 166—173

[5] KARAMARDIAN S. Complementarity Problems Over Cones with Monotone and Pseudomonotone Maps[J]. Journal of Optimization Theory and Applications, 1976, 18(4):445—454

[6] CHAMBERS M L, PONTRYAGIN L S, BOLTYANSKII V G, et al. The Mathematical Theory of Optimal Processes [M]. Oxford:Pergamon Press,1964

[7] PENG H J, GAO Q, ZHANG H W, et al. Parametric Variational Solution of Linear-Quadratic Optimal Control Problems with Control Inequality Constraints[J]. Applied Mathematics and Mechanics,2014,35(9): 1079—1098

[8] BUDAK B M, BERKOVICH E M, SOLOVEVA E N. Difference Approximations in Optimal Control Problems[J]. SIAM Journal on Control,1969, 7(1):18—31

[9] PULULOVA N V. A Pointwise Projected Gradient Method Applied to an Optimal Control Problem[J]. Journal of Computational and Applied Mathematics, 2009,226(2): 331—335

[10] 窦莉峰. 用交替方向法求解离散线性二次最优控制问题[D]. 石家庄:河北工业大学,2015

DOU L F, Solve Discrete Linear-Quadratic Optimal Control Problem by Alternating Direction Method[D]. Shijiazhuang: Hebei University of Technology, 2015(in Chinese)

[11] SOLODOV M V. Convergence Rate Analysis of Iterative Algorithms for Solving Variational Inequality Problems[J]. Math Program,2003(96) 513—528

The First Order Sufficient Conditions for Mayer Type Linear Optimal Control Problems

CHUN Li, WANG Bin

(School of Mathematical Science, Sichuan Normal University,Chengdu 610068, China)

Abstract:The first order sufficient optimality condition for Mayer-type optimal control problems is established under the condition that the cost function is pseudoconvex. The control system is assumed to be linear. In addition, the first order sufficient optimality condition for the corresponding discrete optimal control problem is also proved. As an application, the discrete optimal control problem is equivalently transformed into a finite dimensional variational inequality, and a numerical example is given by the numerical algorithm for pseudo-monotone variational inequality.

Key words:optimal control problem; sufficient condition; variational inequality; pseudo-convex function

责任编辑:李翠薇

引用本文/Cite this paper:

淳黎,王彬. Mayer 型线性最优控制问题的一阶充分条件[J]. 重庆工商大学学报(自然科学版),2020,37(6):56—61

CHUN L, WANG B. The First Order Sufficient Conditions for Mayer Type Linear Optimal Control Problems[J]. Journal of Chongqing Technology and Business University (Natural Science Edition),2020,37(6):56—61