第 30 卷第 9 期
2013 年 9 月

控 制 理 论 与 应 用
Control Theory & Applications

Vol. 30 No. 9
Sep. 2013

DOI: 10.7641/CTA.2013.21101

# 离散时间最优控制——评论动态规划

吴受章†

(西安交通大学, 陕西 西安 710049)

**摘要**: 阐述离散时间最优控制的特点. 对比3种求解离散时间最优控制的解法, 即: 1) 用非线性规划求解离散时间最优控制; 2) 用无约束优化求解离散时间最优控制; 3) 动态规划及其数值解. 1)和2)都适用于多维静态优化, 计算效率较高, 是高级方法. 在名义上, 3)为动态优化. 实际上, 3)为一维分段无约束静态优化, 计算效率较低, 是初级方法. 本文并用数字实例进一步阐明动态规划及其数值解在求解方面较差, 故动态规划及其数值解已失去实用价值. 在求解离散时间最优控制问题方面, 无法与非线性规划求解相匹敌.

**关键词**: 最优控制; 非线性规划; 动态规划

**中图分类号**: TP13　　　**文献标识码**: A

# Discrete-time optimal control— comments on dynamic programming

WU Shou-zhang†

(Xi'an Jiao Tong University, Xi'an Shaanxi 710049, China)

**Abstract:** Characteristics of discrete-time optimal control are investigated. Comparisons of three kinds of methods for solving discrete-time optimal control are given; namely: 1) nonlinear programming to solve discrete-time optimal control; 2) unconstrained optimization to solve discrete-time optimal control; 3) dynamic programming and its numerical solution. Methods 1) and 2) are applicable to multidimensional static optimization, the computation efficiency is high; thus, they are the advanced methods. Although 3) is nominally the dynamic optimization, it is actually the one-dimensional unconstrained piecewise static optimization with low computation efficiency. Thus, it is the elementary method only. Numerical examples illustrate that dynamic programming and its numerical solution are worse in problem solving. Hence, dynamic programming and its numerical solution have lost their practical value, and is unable to compete with the nonlinear programming in solving discrete-time optimal control problems.

**Key words:** optimal control; nonlinear programming; dynamic programming

## 1 引言(Introduction)

离散时间最优控制在军事工业和民用工业中有广泛应用. 动态规划及其数值解法在求解离散时间最优控制的方法中占有主导地位. 在有关最优控制的教材及讲授课程中, 对动态规划和非线性规划而言, 也只介绍动态规划及其数值解法, 但用非线性规划求解离散时间最优控制的方法始终被视为无足轻重, 可有可无, 并已经被人们逐渐淡忘.

从名称上讲, 动态规划是动态优化, 而非线性规划是静态优化. 因此, 动态规划在名称上已经先占有优势. 从来源来讲, 动态规划由控制理论家Bellman提出, 并被誉为最优控制理论发展的里程碑之一, 而非线性规划在控制理论界和控制工程界看来, 仅为外来的普通的数学工具而已. 因此, 动态规划在来源上又占有优势.

当动态规划一经面世, 就以其新颖获得人们首肯, 但人们也看到了它有缺点. 于是人们一方面举双手供奉它, 另一方面又不断声称要改进它. 遗憾的是, 经历了半个多世纪, 动态规划还是面目依旧, 改进不了. 对此, 人们却已习以为常.

本文首先阐述离散时间最优控制的特点, 然后, 对比3种解法, 即: 1) 用非线性规划求解离散时间最优控制; 2) 动态规划及其数值解法; 3) 用无约束优化求解离散时间最优控制. 其中, 无约束优化仍然属于非线性规划范畴, 却因为它是无约束, 故单列一项. 继而, 用数字实例进一步阐明用非线性规划求解离散时间最优控制, 及用无约束优化求解离散时间最优控制, 都比较好; 而动态规划及其数值解法比较差. 最后得出结论.

## 2 离散时间最优控制的特点(Discrete time optimal control characteristics)

离散时间最优控制的提法为

收稿日期: 2012−10−25; 收修改稿日期: 2013−03−28.
†通信作者. E-mail: wsz_1@xjtu.edu.cn.
基金项目: 国家自然科学基金资助项目(60974101, 61221063).

$$\min_{u(k)} J = \theta(x(N)) + \sum_{k=0}^{N-1} F(x(k), u(k)), \quad (1)$$
$$\text{s.t. } x(k+1) = f(x(k), u(k)), x(0) = x_0,$$

式中: $f$为非时变系统, 目标函数亦为非时变的; $x \in \mathbb{R}^1$, $u \in \mathbb{R}^1$, $f \in \mathbb{R}^1$, $k$为离散时刻, $N$为段数. 注意到式(1)实质上是静态优化, 它也具有分段静态优化的特点, 但有约束.

式(1)中, 末两项记为

$$\begin{aligned} &\min G_{N-1} = \\ &\min\{F(x(N-1), u(N-1)) + \\ &\theta[f(x(N-1), u(N-1))]\}, \end{aligned} \quad (2)$$

式中$N-1$段之前的一切信息集中反映于$x(N-1)$之中, 故可令

$$\frac{\partial}{\partial u(N-1)} G_{N-1} = 0, \quad (3)$$

可得

$$P_{N-1}(x^*(N-1), u^*(N-1)) = 0, \quad (4)$$

式中$*$记最优值(以下, 为书写方便, $*$被去掉, 但需要时又被添上). 式(4)为二维曲线, 满足式(4)的$x^*(N-1)$和$u^*(N-1)$有无穷多组解.

若式(4)可显化, 代入$G_{N-1}$中, 得

$$\min G_{N-1} = g_{N-1}(x^*(N-1)).$$

式(1)中, 末3项记为

$$\begin{aligned} \min G_{N-2} = \min\{&F(x(N-2), u(N-2)) + \\ &g_{N-1}[f(x(N-2), u(N-2))]\}. \end{aligned}$$

仿式(3)的做法, 可得

$$P_{N-2}(x^*(N-2), u^*(N-2)) = 0. \quad (5)$$

类此,

$$P_{N-3}(x^*(N-3), u^*(N-3)) = 0, \quad (6)$$
$$\vdots$$
$$P_1(x^*(1), u^*(1)) = 0, \quad (7)$$
$$P_0(x^*(0), u^*(0)) = 0. \quad (8)$$

综合之, 存在一个特殊的并集

$$\bigcup_{k=0}^{N-1} \{P_k(x^*(k), u^*(k)) = 0\}. \quad (9)$$

式(4)–(9)都是对分段目标函数求min的结果, 但分段目标函数的范围随$k$增大, 式(8)已成为对总目标函数求min的结果. 即式(8)及式(9)都考虑了总目标.

式(9)所示特殊的并集有3个特点,

1) 由$N$个集合构成有限并集, $k = 0, 1, 2, \cdots, N-1$.

2) $N$个集合中的每一个集合, 都含有无限多元素.

3) 满足总目标为min的$x^*(k), u^*(k)$存在于该特殊的并集中.

式(9)可用来鉴别对式(1)的解法是否可实现.

## 3 评论动态规划(Comments on dynamic programming)

对式(1)的解法, 有如下3种是可实现的: 1) 非线性规划; 2) 动态规划及其数值解法; 3) 无约束优化(因为它是无约束, 故单列一项). 分述于下:

1) 非线性规划.

用非线性规划求解离散时间最优控制早已有之, 虽然它和动态规划是同一时代的产物, 但人们重视动态规划, 不重视前者. 人们曾认为: 前者即使对一维状态方程, 也要把所有$x(k), u(k)$都作为决策变量, 所以维数增大了, 这是个严重缺点. 动态规划面对一维状态方程, 一维就是一维, 所以, 动态规划更受欢迎. 另一方面, 当年, 前者技术还不成熟, 遇到了一些困难:

a) 缺乏优秀的非线性规划(并非一切非线性规划都适用);

b) 即使是线性状态方程, 把所有$x(k), u(k)$都作为决策变量后, 会出现稀疏矩阵约束, 则应如何序列修正Lagrange乘子向量, 并提高计算精度;

c) 没有人机界面良好的软件.

但是, 现在前者已克服了这些困难. 还有, 从表面上看, 把所有$x(k), u(k)$都作为决策变量, 的确使维数增大了, 是个严重缺点; 但面对多维或高维状态方程, 无非是多增加了些决策变量, 事实上并未产生所谓"维数灾难", 也并未使求解的难度发生质的变化, 这却是个值得庆幸的优点. 况且, 非线性规划又特别擅长于求解多维或高维问题. 此外, 非线性规划还特别擅长于考虑有界约束, 等式及不等式约束. 第4节将进一步看到非线性规划比动态规划在求解方面更好.

从$k = 0$开始, 用非线性规划求解约束优化问题, 在各段上同时优化下一步. 效率较高.

2) 动态规划及其数值解法.

用动态规划求解离散时间最优控制, 只能用于一维状态方程和简单的目标函数. 目标函数从末段开始反向扫掠; 状态方程从$k = 0$ 开始正向扫掠. 遗憾的是:

a) 动态规划采用一维分段无约束优化, 其计算效率较低;

b) 名为动态规划, 实为一维分段无约束优化, 并且仅为普通的求偏导;

c) 同样是化为无约束优化问题, 逐段将状态方程代入目标函数, 不如一次性将状态方程完全代入目标函数;

d) 状态方程正向扫掠仅为解代数方程. 第4节将进一步看到与非线性规划相比, 动态规划在求解方面较差.

为了解决动态规划的计算机求解, 传统采用状态空间网格化(量化)的一种数值解法[1], 此法能考虑各种复杂的约束. 其缺点为:

a) 必须预知状态方程解的分布, 否则量化是盲目的, 并且, 若量化范围设置不当, 会导致无法计算;

b) 过粗的量化, 使计算不准确, 而过细的量化, 又使得难以计算;

c) 不能用于多维或高维状态方程(只能用于一维状态方程);

d) 用有级的状态变量取代无级的状态变量, 使计算精度降低;

e) 所谓“维数灾难”, 正是由文[1]自己把状态空间网格化造成的.

第4节将进一步看到与非线性规划相比, 动态规划及其数值解法在求解方面较差.

3) 无约束优化.

一次性将状态方程完全代入目标函数, 用无约束优化求解,效率较高. 但不能考虑有界约束, 其他等式及不等式约束. 第4节将进一步看到无约束优化比动态规划在求解方面更好.

第2节已说明, 式(1)所示离散时间最优控制问题实质上是静态优化问题, 本节说明非线性规划和无约束优化都是静态优化方法; 动态规划在名义上是动态优化, 实为一维分段无约束静态优化方法. 同为静态优化方法, 动态规划的效率却较低. 至此人们可能才会领悟到, 动态规划是一维分段无约束静态优化方法, 故不可能从动态优化的角度, 并用动态优化方法去改进的, 半个多世纪的历程证实了这一点. 唯一的出路是采用高一级的方法, 才能获得改进.

## 4 数字实例(Numerical examples)

以下数字实例仅用于阐明算法的使用结果:

**例 1**

$$\min_{u(k)} J = x^2(3) + \sum_{k=0}^{2}(x^2(k), u^2(k)),$$
$$\text{s.t. } x(k+1) = x(k) + u(k),\ x(0) = 1. \tag{10}$$

采用MATLAB优化工具箱的约束非线性优化求解器fmincon, 并分别用算法interior-point, 或active-set, 或sqp, 三者都获得同样结果, 见表1.

表 1 式(10)的最优解

Table 1 The optimal solution of Eq.(10)

| $x(0)$ | $x(1)$ | $x(2)$ | $x(3)$ | $u(0)$ | $u(1)$ | $u(2)$ |
|---|---|---|---|---|---|---|
| 1.0000 | 0.3846 | 0.1538 | 0.0769 | −0.6154 | −0.2308 | −0.0769 |

采用MATLAB全局优化工具箱的遗传算法求解器ga, 或将状态方程完全代入目标函数后, 采用MATLAB优化工具箱的无约束非线性优化求解器fminunc, 或采用动态规划(手算或用MATLAB的符号数学工具箱仿效手算), 都可获得表 1 同样的结果. 但ga的迭代次数多, 精度差些.

例1是各种算法都能计算并相互比较的起点.

**例 2**

$$\min_{u(k)} J = 0.5\sum_{k=0}^{9} u^2(k),$$
$$\text{s.t. } x(k+1) = x(k) + u(k),$$
$$x(0) = 1,\ x(10) = 0. \tag{11}$$

采用MATLAB优化工具箱的约束非线性优化求解器fmincon, 并分别用算法interior-point, 或active-set, 或sqp, 三者都获得同样结果, 见表2.

表 2 式(11)的最优解

Table 2 The optimal solution of Eq.(11)

| $x(0)$ | $x(1)$ | $x(2)$ | $x(3)$ | $x(4)$ | $x(5)$ | $x(6)$ | $x(7)$ | $x(8)$ | $x(9)$ | $x(10)$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1.0 | 0.9 | 0.8 | 0.7 | 0.6 | 0.5 | 0.4 | 0.3 | 0.2 | 0.1 | 0.0 |
| $u(0)$ | $u(1)$ | $u(2)$ | $u(3)$ | $u(4)$ | $u(5)$ | $u(6)$ | $u(7)$ | $u(8)$ | $u(9)$ | $u(10)$ |
| −0.1 | −0.1 | −0.1 | −0.1 | −0.1 | −0.1 | −0.1 | −0.1 | −0.1 | −0.1 | — |

式(11)的解析解为

$$u(k) = -1/10, \tag{12}$$

$$x(k) = 1 - k/10. \tag{13}$$

表2与式(12)–(13)所示解析解的计算结果完全一致.

采用MATLAB 全局优化工具箱的遗传算法求解器ga, 或将状态方程完全代入目标函数后, 采用MATLAB 优化工具箱的无约束非线性优化求解器fminunc; 都可获得表2同样的结果. 但ga的迭代次数多, 精度差些.

但是, 采用动态规划的数值解法, 按惯例, 若将状态变量和控制变量分别量化为

$$X = \{-6, -5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5, 6\} \tag{14}$$

$$U = \{-6, -5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5, 6\}. \tag{15}$$

由于在计算前, 未能预知解的分布, 量化范围设置不当, 导致无法计算. 换言之, 动态规划的数值解法的使用过程非常不爽(特别是还要遭受没有通用程序之苦).

**例 3**

$$\begin{aligned}&\min_{u(k)} J = \sum_{k=0}^{4}(x^{\mathrm{T}}(k)x(k) + u^{\mathrm{T}}(k)u(k)),\\ &\text{s.t. } x(k+1) = Ax(k) + Bu(k),\ x(0) = 1,\\ &\quad x_1^2(5) + x_2^2(5) + x_3^2(5) \leqslant 1,\end{aligned} \tag{16}$$

式中: $x \in \mathbb{R}^3,\ u \in \mathbb{R}^2$,

$$\begin{cases} A = \begin{bmatrix} 2.8971 & 0.2530 & -0.9024 \\ 0.0 & 0.0548 & 0.9271 \\ 0.4882 & -4.7601 & 0.8607 \end{bmatrix}, \\ B = \begin{bmatrix} 0.0380 & 0.0 \\ 1.0 & 0.0808 \\ 0.0 & 0.2298 \end{bmatrix}. \end{cases} \tag{17}$$

采用MATLAB优化工具箱的约束非线性优化求解器fmincon, 并分别用算法interior-point, 或active-set, 或sqp, 三者都获得同样结果, 见表3. 动态规划及其数值解法对此例已无能为力.

例3的背景是制导飞行器到达终端函数. 改写式(16)后, 还可以在敌方的拦截区内产生角隅条件, 避开拦截.

表 3 式(16)的最优解

Table 3 The optimal solution of Eq.(16)

| $u_1(0)$ | $u_2(0)$ | $u_1(1)$ | $u_2(1)$ | $u_1(2)$ | $u_2(2)$ | $u_1(3)$ | $u_2(3)$ | $u_1(4)$ | $u_2(4)$ |
|---|---|---|---|---|---|---|---|---|---|
| −4.9621 | 16.2299 | 1.5332 | −0.6119 | −11.6676 | −1.4865 | −5.5505 | −0.0706 | −0.7517 | 0.1512 |
| $x_1(0)$ | $x_2(0)$ | $x_3(0)$ | $x_1(1)$ | $x_2(1)$ | $x_3(1)$ | $x_1(2)$ | $x_2(2)$ | $x_3(2)$ | — |
| 1.0000 | 1.0000 | 1.0000 | 2.0591 | −2.6688 | 0.3184 | 5.0612 | 1.6328 | 13.8427 | — |
| $x_1(3)$ | $x_2(3)$ | $x_3(3)$ | $x_1(4)$ | $x_2(4)$ | $x_3(4)$ | $x_1(5)$ | $x_2(5)$ | $x_3(5)$ | — |
| 2.1409 | 1.1354 | 6.2716 | 0.6193 | 0.3204 | 1.0224 | 0.9242 | 0.2260 | −0.3080 | — |

**例 4**

$$\begin{aligned}&\min_{u(k)} J = x^{\mathrm{T}}(5)x(5) + \sum_{k=0}^{4}(x^{\mathrm{T}}(k)x(k) + u^{\mathrm{T}}(k)u(k)),\\ &\text{s.t. } x(k+1) = Ax(k) + Bu(k),\ x(0) = 1,\ x(k) \geqslant 0,\end{aligned} \tag{18}$$

式中: $x \in \mathbb{R}^3,\ u \in \mathbb{R}^2$,

$$A = \begin{bmatrix} 2.8971 & 0.2530 & -0.9024 \\ 0.0 & 0.0548 & 0.9271 \\ 0.4882 & -4.7601 & 0.8607 \end{bmatrix},\ B = \begin{bmatrix} 0.0380 & 0.0 \\ 1.0 & 0.0808 \\ 0.0 & 0.2298 \end{bmatrix}. \tag{19}$$

表 4 式(18)的最优解

Table 4 The optimal solution of Eq.(18)

| $u_1(0)$ | $u_2(0)$ | $u_1(1)$ | $u_2(1)$ | $u_1(2)$ | $u_2(2)$ | $u_1(3)$ | $u_2(3)$ | $u_1(4)$ | $u_2(4)$ |
|---|---|---|---|---|---|---|---|---|---|
| −3.4423 | 30.4500 | −3.8651 | 6.6877 | −4.4902 | 0.0007 | −5.0024 | −0.4280 | −1.4713 | 0.0238 |
| zc $x_1(0)$ | $x_2(0)$ | $x_3(0)$ | $x_1(1)$ | $x_2(1)$ | $x_3(1)$ | $x_1(2)$ | $x_2(2)$ | $x_3(2)$ | — |
| 1.0000 | 1.0000 | 1.0000 | 2.1169 | 0.0000 | 3.5862 | 2.7498 | 0.0000 | 5.6569 | — |
| $x_1(3)$ | $x_2(3)$ | $x_3(3)$ | $x_1(4)$ | $x_2(4)$ | $x_3(4)$ | $x_1(5)$ | $x_2(5)$ | $x_3(5)$ | — |
| 2.6910 | 0.7544 | 6.2115 | 2.1915 | 0.7631 | 2.9708 | 3.8052 | 1.3266 | 0.0000 | — |

采用MATLAB优化工具箱的约束非线性优化求解器fmincon, 并分别用算法interior-point, 或active-set, 或sqp, 三者都获得同样结果, 见表4. 动态规划及其数值解法对此例已无能为力.

类似于例4的背景, 改写式(18)后, 还可以制导飞行器紧贴水面飞行. 还可以制导飞行器低空或超低空飞行, 以避开敌方雷达跟踪.

以上的数字实例进一步说明动态规划及其数值解法在求解方面较差.

## 5　结论(Conclusions)

在求解离散时间最优控制时, 不宜采用动态规划及其数值解法(网格化); 宜采用非线性规划或无约束优化.

一维优化是初级方法, 多维优化是高级方法. 动态规划及其数值解法属于前者, 非线性规划属于后者.

实际上, 用非线性规划求解离散时间最优控制可取代动态规划及其数值解.

**致谢**　西安交通大学系统所和高峰老师.

## 参考文献(References):

[1] LARSON R E, CASTI J L. *Principles of Dynamic Programming, Part II* [M]. New York: Marcel Dekker, Inc, 1982: 233 – 345.

## 作者简介:

吴受章　(1934—), 男, 1957年毕业于上海交通大学电机系, 曾在上海交通大学电机系、西安交通大学电机系工作, 后在西安交通大学系统所工作; 曾讲授电磁与电子控制装置、自动控制原理、最优控制、自适应控制等课程; 曾编著应用最优控制、最优控制理论与应用等书籍; 曾研究磁性分频器、同步发电机励磁控制、模型降阶与控制、大气污染控制、自适应控制、柔性制造系统的建模与控制等, Email: wsz_1@xjtu.edu.cn.