

求解一维椭圆最优控制问题的离散序列 二次哈密顿算法^{*}

向琴琴 陈浩[†]

(重庆师范大学 数学科学学院, 重庆 401331)

摘要 近期, Breitenbach 和 Borzi 在“先优化再离散”框架下提出了一类求解椭圆最优控制问题的序列二次哈密顿方法. 他们证明了该迭代方法在连续空间情形下的单调收敛性. 然而, 在“先离散再优化”框架下, 该迭代方法在离散情形下的收敛性质尚未被解决. 本文探讨了一维椭圆最优控制问题的一类中心差分离散格式, 并证明了其能保持序列二次哈密顿迭代的单调收敛性. 数值实验表明了该方法是有效且收敛的.

关键词 椭圆最优控制问题, 序列二次哈密顿方法, 差分格式, 收敛性

中图分类号: O241.81

文献标志码: A

Discrete Sequential Quadratic Hamiltonian Algorithm for One Dimensional Elliptic Optimal Control Problems^{*}

Xiang Qinqin Chen Hao[†]

(College of Mathematics Science, Chongqing Normal University, Chongqing 401331, China)

Abstract Recently, Breitenbach and Borzi have proposed a sequential quadratic Hamiltonian method for solving elliptic optimal control problems in the optimize-then-discretize framework. They prove the monotonic convergence of the algorithm in the continuous space case. However, the properties of the discrete version of the iterative procedure have not been yet tackled with in the discretize-then-optimize framework. In this paper we present a center difference scheme for one-dimensional elliptic optimal control problems and prove that the scheme preserves the monotonic properties of the sequential quadratic Hamiltonian method. Numerical experiments show that the proposed algorithm is effective and convergent.

Key words elliptic optimal control problems, sequential quadratic Hamiltonian method, difference scheme, convergence

引言

偏微分方程(PDE)最优控制问题的应用广泛, 覆盖了医学^[1]、生态学^[2]、工程学^[3]等诸多领域, 具有重要的实际意义. 通过优化控制变量, PDE 最优控制问题能够显著提高效率、降低成本、改善性能, 从而推动各行业技术的进步和创新. 最优控制理论

的研究方法多种多样, 主要包括 Pontryagin 等^[4]提出的 Pontryagin 极大值原理, Bellman^[5]创立的动态规划方法以及 Kalman^[6]发展出的线性二次系统的最优控制理论. 这些方法为最优控制理论的发展奠定了坚实的基础.

在 PDE 约束优化问题的数值求解中, 根据问题结构和计算需求的不同, 目前主要存在两种求解

2025-02-25 收到第 1 稿, 2025-05-09 收到修改稿.

^{*} 重庆市自然科学基金项目资助(cstc2021jcyj-msxmX0034), Natural Science Foundation Chongqing Project (cstc2021jcyj-msxmX0034).

[†] 通信作者 E-mail: hch@cqu.edu.cn

策略:先离散再优化和先优化再离散.先离散再优化是先对状态变量和控制变量进行参数化处理,将最优控制问题转化为一系列有限维的优化问题,再选择合适的优化算法进行数值求解.先离散再优化在实践中表现出很强的鲁棒性,可以参考文献[7].但是随着离散的空间和时间步长不断减小,计算量呈非线性快速增长.先优化再离散是先连续空间中推导出最优控制问题的最优性条件(如 Pontryagin 极大值原理),得到连续的最优性系统(包括状态方程和伴随方程),再对其进行离散求解.

PDE 约束优化问题是一个无穷维空间上的优化问题,涉及函数空间的优化算法和离散方法等方面.随着研究的不断深入,求解 PDE 约束优化问题的方法不断丰富,常用的优化算法包括原对偶积极集方法^[8,9]、序列二次规划方法(SQP)^[10,11]、增广 Lagrange 方法^[12]、交替方向乘子算法(ADMM)^[13]等;常用的离散方法包括有限元方法^[14]、谱方法^[15]、有限差分法^[16]等.

针对椭圆最优控制问题的数值求解,研究者们基于不同特征提出了多种有效的离散方法.例如 Dai 和 Chen^[17]使用有限元方法处理非线性椭圆方程约束的最优控制问题,从理论上严格证明了该方法具有超收敛特性.在控制与状态双重约束情形下,Chen 等^[18]研究了椭圆最优控制问题的 Galerkin 谱方法,不仅推导了先验误差估计,还推导了后验误差估计.Chen 等^[19]研究了将谱方法推广至椭圆及抛物最优控制问题,并给出了该方法的误差估计结果.

基于离散最优控制系统的特点,选择合适的优化算法至关重要.Li 等^[20]将共轭梯度法与有限体积元素法结合,用于求解线性椭圆最优控制问题,并给出了先验误差估计.Song 等^[21]提出一种具有非精确结构的交替方向乘子算法,用于求解带 L^1 范数控制代价的椭圆最优控制问题,不仅证明了算法的全局收敛性,还分析了其收敛速率.Zhang 等^[22]使用了带线搜索条件的共轭梯度算法,用于求解线性椭圆最优控制问题,获得了包括全局收敛性和线性收敛速率在内的完整理论结果.

近期,Borzi 等^[23,24]在“先优化再离散”框架下提出了一类求解最优控制问题的序列二次哈密顿(SQH)算法,该方法基于 Pontryagin 极大值原理,且其特别适用于求解非光滑、非凸最优控制问题.

SQH 方法通过逐点优化增广哈密顿函数来实现求解,具体过程为:该方法首先引入初始控制变量,并基于此计算出初始状态变量,然后进入迭代循环阶段,依次更新伴随变量、控制变量和状态变量,直到满足收敛条件.SQH 迭代算法在求解椭圆最优控制问题时表现出了优良的鲁棒性和高效性,Breitenbach 和 Borzi 已证明 SQH 迭代方法在连续空间情形下是单调收敛性的,但其在“先离散后优化”框架下的对应算法及单调收敛性质尚未被讨论.

鉴于此,本文的主要目的是探讨在“先离散再优化”的框架下 SQH 算法的单调收敛性.我们主要考虑一维椭圆最优控制问题,在“先离散再优化”框架下,针对一类中心差分离散格式,我们得到了离散的 SQH 算法,并证明了该方法的单调收敛性.数值实验验证了该方法的单调收敛性.

1 椭圆最优控制问题

本文考虑如下的最优控制问题:

$$\min_{u \in U_{\text{ad}}} J(y, u) = \min_{u \in U_{\text{ad}}} \int_a^b l[x, y(x), u(x)] dx \quad (1)$$

满足状态方程约束

$$\begin{cases} -y'' = f[x, y(x), u(x)], & x \in (a, b) \\ y'(a) - \alpha y(a) = g_1 \\ y'(b) - \beta y(b) = g_2 \end{cases} \quad (2)$$

式中, $y: [a, b] \rightarrow R$ 是状态变量,表示系统的状态; $u: [a, b] \rightarrow R$ 是控制变量,通过影响系统来达到最优化目标; $l: [a, b] \times R \times R \rightarrow R$ 是运行成本; $f: [a, b] \times R \times R \rightarrow R$; $J(y, u)$ 是目标泛函; $\alpha, \beta, g_1, g_2 \in R$ 且 $\alpha \geq 0, \beta \leq 0$. 这里控制变量的约束集为:

$$U_{\text{ad}} = \{u(x) \in L^2(a, b); u(x) \in K_{\text{ad}}, \forall x \in [a, b]\} \quad (3)$$

式中, $L^2(a, b)$ 是 $[a, b]$ 上的平方可积函数空间, K_{ad} 是 R 中的紧致子集.

对于上述最优控制问题,我们定义哈密顿函数:

$$H(x, y, u, p) = pf(x, y, u) + l(x, y, u) \quad (4)$$

式中, $p: [a, b] \rightarrow R$ 为伴随变量,其代表了对应于状态变量 y 的拉格朗日乘子.

由 Pontryagin 极大值原理^[25]可知,最优控制问题(1)~(3)的一阶必要条件为:

$$\begin{aligned} -y'' &= H_p(y, u, p) = f(x, y, u) \\ y'(a) - \alpha y(a) &= g_1 \\ y'(b) - \beta y(b) &= g_2 \end{aligned} \quad (5)$$

$$\begin{aligned} -p'' &= H_y(x, y, u, p) = f_y(x, y, u)p + l_y(x, y, u) \\ p'(a) - \alpha p(a) &= 0 \\ p'(b) - \beta p(b) &= 0 \end{aligned} \quad (6)$$

$$\begin{aligned} u(x) &= \arg \min_{v \in K_{ad}} H[x, y(x), v, p(x)], \\ \forall x &\in [a, b] \end{aligned} \quad (7)$$

为了数值实现(5)~(7), Borzi 等引入了增广哈密顿函数:

$$\begin{aligned} H_\epsilon(x, y, u, v, p) &= H(x, y, u, p) + \\ &\epsilon(u - v)^2 \end{aligned} \quad (8)$$

式中, $\epsilon > 0$ 为正则化参数, 并设计了一类单调收敛的 SQH 方法(参见文献[23]). 该 SQH 方法在“先离散再优化”框架下是否仍能保持单调收敛性质是一个有意义且尚未明确解决的问题. 本文将构造求解一维椭圆最优控制问题的一类基于二阶中心差分格式的离散序列二次哈密顿算法, 并探究其单调收敛性.

2 空间离散格式

我们将区间 $[a, b]$ 等距划分, 空间网格为 $a = x_0 < x_1 < \dots < x_N = b$, 空间步长 $h = (b - a)/N$. 考虑状态方程(2)的如下二阶中心差分格式:

$$\begin{aligned} -\frac{y_{n-1} - 2y_n + y_{n+1}}{h^2} &= f(x_n, y_n, u_n), \\ n &= 1, 2, \dots, N-1 \end{aligned} \quad (9)$$

式中, y_n 和 u_n 分别是 $y(x_n)$ 和 $u(x_n)$ 的近似解.

对于边界处的离散, 我们引入虚拟变量 y_{-1} 和 y_{N+1} , 对微分方程在 $x=a, b$ 处利用中心差分离散, 对边界条件利用一阶导的中心差分离散, 消去虚拟变量 y_{-1} 和 y_{N+1} , 最终得到边界点的差分格式:

$$\frac{(1 + h\alpha)y_0 - y_1}{h^2} = -\frac{g_1}{h} + \frac{1}{2}f(x_0, y_0, u_0) \quad (10)$$

$$\frac{-y_{N-1} + (1 - h\beta)y_N}{h^2} = \frac{g_2}{h} + \frac{1}{2}f(x_N, y_N, u_N) \quad (11)$$

我们利用复合梯形公式 $\frac{h}{2}l(x_0, y_0, u_0) + h \sum_{n=1}^{N-1} l(x_n, y_n, u_n) + \frac{h}{2}l(x_N, y_N, u_N)$ 来近似目标泛函中的积分项 $\int_a^b l[x, y(x), u(x)]dx$.

应用上述离散格式, 我们可以定义如下离散最优控制问题:

$$\begin{aligned} \min J_N(\mathbf{y}, \mathbf{u}) &= \frac{h}{2}l(x_0, y_0, u_0) + h \sum_{n=1}^{N-1} l(x_n, \\ &y_n, u_n) + \frac{h}{2}l(x_N, y_N, u_N) \end{aligned} \quad (12)$$

满足

$$\begin{cases} \frac{(1 + h\alpha)y_0 - y_1}{h^2} = -\frac{g_1}{h} + \frac{1}{2}f(x_0, y_0, u_0) \\ \frac{-y_{n-1} + 2y_n - y_{n+1}}{h^2} = f(x_n, y_n, u_n), n = 1, 2, \dots, N-1 \\ \frac{-y_{N-1} + (1 - h\beta)y_N}{h^2} = \frac{g_2}{h} + \frac{1}{2}f(x_N, y_N, u_N) \end{cases} \quad (13)$$

式中, $\mathbf{y} = (y_0, y_1, \dots, y_N)^T$, $\mathbf{u} = (u_0, u_1, \dots, u_N)^T$.

上述优化问题相应的拉格朗日泛函为:

$$\begin{aligned} L(\mathbf{y}, \mathbf{u}, \mathbf{p}) &= \frac{h}{2}l(x_0, y_0, u_0) + \\ &h \sum_{n=1}^{N-1} l(x_n, y_n, u_n) + \frac{h}{2}l(x_N, y_N, u_N) + \\ &\frac{h}{2}p_0 \left[\frac{(-2 - 2h\alpha)y_0 + 2y_1}{h^2} - \frac{2g_1}{h} + f(x_0, y_0, u_0) \right] + \\ &h \sum_{n=1}^{N-1} p_n \left[\frac{y_{n-1} - 2y_n + y_{n+1}}{h^2} + f(x_n, y_n, u_n) \right] + \\ &\frac{h}{2}p_N \left[\frac{2y_{N-1} + (-2 + 2h\beta)y_N}{h^2} + \right. \\ &\left. \frac{2g_2}{h} + f(x_N, y_N, u_N) \right] \end{aligned} \quad (14)$$

式中, $\mathbf{p} = (p_0, p_1, \dots, p_N)^T$ 为(13)的拉格朗日乘子.

离散优化问题(12)、(13)的一阶最优性条件可通过如下方式获得:

$$\begin{aligned} \frac{\partial L(\mathbf{y}, \mathbf{u}, \mathbf{p})}{\partial p_n} &= 0, \quad \frac{\partial L(\mathbf{y}, \mathbf{u}, \mathbf{p})}{\partial y_n} = 0, \\ n &= 0, 1, \dots, N \end{aligned} \quad (15)$$

由此我们可以得到离散的伴随方程:

$$\begin{aligned} \frac{(1 + h\alpha)p_0 - p_1}{h^2} &= \frac{1}{2}f_y(x_0, y_0, u_0)p_0 + \\ &\frac{1}{2}l_y(x_0, y_0, u_0) \end{aligned} \quad (16)$$

$$\begin{aligned} \frac{-p_{n-1} + 2p_n - p_{n+1}}{h^2} &= f_y(x_n, y_n, u_n)p_n + \\ &l_y(x_n, y_n, u_n), \quad n = 1, 2, \dots, N-1 \\ \frac{-p_{N-1} + (1 - h\beta)p_N}{h^2} &= \frac{1}{2}f_y(x_N, y_N, u_N)p_N + \end{aligned} \quad (17)$$

$$\frac{1}{2}l_y(x_N, y_N, u_N) \tag{18}$$

对比状态方程的差分离散,我们不难发现(16)~(18)为连续伴随方程(6)的中心差分离散,即有 p_n 为 $p(x_n)$ 的近似解.因此,我们先离散后优化所得到的离散系统为先优化所得到的 一阶连续最优性条件(5)、(6)的相容逼近,即优化与离散是相容的.

利用哈密顿函数(4),离散状态方程与离散伴随方程可以重写为哈密顿结构,进一步可获得离散的一阶最优性条件:

$$\frac{(1+h\alpha)y_0-y_1}{h^2}=-\frac{g_1}{h}+\frac{1}{2}H_p(x_0,y_0,u_0,p_0) \tag{19}$$

$$\frac{-y_{n-1}+2y_n-y_{n+1}}{h^2}=H_p(x_n,y_n,u_n,p_n), \tag{20}$$
$$n=1,2,\cdots,N-1$$

$$\frac{-y_{N-1}+(1-h\beta)y_N}{h^2}=\frac{g_2}{h}+\frac{1}{2}H_p(x_N,y_N,u_N,p_N) \tag{21}$$

$$\frac{(1+h\alpha)p_0-p_1}{h^2}=\frac{1}{2}H_y(x_0,y_0,u_0,p_0) \tag{22}$$

$$\frac{-p_{n-1}+2p_n-p_{n+1}}{h^2}=H_y(x_n,y_n,u_n,p_n), \tag{23}$$
$$n=1,2,\cdots,N-1$$

$$\frac{-p_{N-1}+(1-h\beta)p_N}{h^2}=\frac{1}{2}H_y(x_N,y_N,u_N,p_N) \tag{24}$$

$$u_n=\operatorname{argmin}_{v\in K_{\text{ad}}}H(x_n,y_n,v,p_n),n=0,1,\cdots,N \tag{25}$$

令

$$\mathbf{A}=\begin{bmatrix} 1+h\alpha & -1 & & \\ -1 & 2 & -1 & \\ & & \ddots & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 1-h\beta \end{bmatrix},$$

$$\mathbf{F}(\mathbf{y},\mathbf{u})=\begin{bmatrix} -\frac{g_1}{h}+\frac{1}{2}f(x_0,y_0,u_0) \\ f(x_1,y_1,u_1) \\ \vdots \\ f(x_{N-1},y_{N-1},u_{N-1}) \\ \frac{g_2}{h}+\frac{1}{2}f(x_N,y_N,u_N) \end{bmatrix},$$

$$\mathbf{B}(\mathbf{y},\mathbf{u})=\operatorname{diag}[\frac{1}{2}f_y(x_0,y_0,u_0),f_y(x_1,y_1,u_1),\cdots,f_y(x_{N-1},y_{N-1},u_{N-1}),\frac{1}{2}f_y(x_N,y_N,u_N)],$$

$$\mathbf{D}(\mathbf{y},\mathbf{u})=[\frac{1}{2}l_y(x_0,y_0,u_0),l_y(x_1,y_1,u_1),\cdots,l_y(x_{N-1},y_{N-1},u_{N-1}),\frac{1}{2}l_y(x_N,y_N,u_N)]^T.$$

则式(19)~(21)可写为矩阵形式

$$\mathbf{A}\mathbf{y}=h^2\mathbf{F}(\mathbf{y},\mathbf{u}) \tag{26}$$

式(22)~(24)可写为

$$\mathbf{A}\mathbf{p}=h^2\mathbf{B}(\mathbf{y},\mathbf{u})\mathbf{p}+h^2\mathbf{D}(\mathbf{y},\mathbf{u}) \tag{27}$$

令

$$\omega_n=\begin{cases} \frac{1}{2}, & n=0,N, \\ 1, & n=1,\cdots,N-1, \end{cases}$$

我们定义如下范数

$$\|\mathbf{y}\|_{\omega}^2=h\langle\mathbf{y},\mathbf{y}\rangle_{\omega}=h\sum_{n=0}^N\omega_ny_n^2 \tag{28}$$

为了实现离散的一阶最优性条件(19)~(25),我们设计了离散的 SQH 算法如表 1 所示.

表 1 离散的 SQH 算法
Table 1 Discrete SQH algorithm

算法 1 离散 SQH 方法求解一维椭圆最优控制问题
输入:初始迭代值 $\mathbf{u}^0$,最大迭代次数 $k_{\max}$,容许误差 $\kappa>0,\epsilon>0,\sigma>0,\eta>0$,且 $\zeta\in(0,1)$,设置 $\tau>\kappa,k=0$.
1) 计算离散状态方程的解 $\mathbf{y}^0$, $\mathbf{A}\mathbf{y}^0=h^2\mathbf{F}(\mathbf{y}^0,\mathbf{u}^0)$. while($k<k_{\max}$ & & $\tau>\kappa$) do
2) 计算离散伴随方程的解 $\mathbf{p}^k$, $\mathbf{A}\mathbf{p}^k=h^2\mathbf{B}(\mathbf{y}^k,\mathbf{u}^k)\mathbf{p}^k+h^2\mathbf{D}(\mathbf{y}^k,\mathbf{u}^k)$.
3) 确定 $\mathbf{u}_n^{k+1}$ 以满足以下最优化问题, $H_{\epsilon}(x_n,y_n^k,u_n^{k+1},u_n^k,p_n^k)=\min_{v\in K_{\text{ad}}}H_{\epsilon}(x_n,y_n^k,v,u_n^k,p_n^k),n=0,1,\cdots,N$.
4) 计算离散状态方程的解 $\mathbf{y}^{k+1}$, $\mathbf{A}\mathbf{y}^{k+1}=h^2\mathbf{F}(\mathbf{y}^{k+1},\mathbf{u}^{k+1})$.
5) 计算 $\tau=\ \mathbf{u}^{k+1}-\mathbf{u}^k\ _{\omega}^2$.
6) 如果 $J_N(\mathbf{y}^{k+1},\mathbf{u}^{k+1})-J_N(\mathbf{y}^k,\mathbf{u}^k)>-\eta\tau$,令 $\epsilon=\sigma\epsilon$,回到第三步. 如果 $J_N(\mathbf{y}^{k+1},\mathbf{u}^{k+1})-J_N(\mathbf{y}^k,\mathbf{u}^k)\leqslant-\eta\tau$,令 $\epsilon=\zeta\epsilon$,进入下一步.
7) $k=k+1$.
end while

3 收敛性分析

本节主要给出离散 SQH 方法的收敛性分析,我们假设最优控制问题(1)~(3)满足如下的条件.

假设 1 在 f 和 l 各自定义域中包含的任何紧致子集中, 存在 Lipschitz 常数 $c_1, c_2 > 0$, 常数 $c_3, c_4 > 0$ 使得

$$|f_y(x, y_1, u_1) - f_y(x, y_2, u_2)| \leq c_1(|y_1 - y_2| + |u_1 - u_2|) \quad (29)$$

$$|l_y(x, y_1, u_1) - l_y(x, y_2, u_2)| \leq c_2(|y_1 - y_2| + |u_1 - u_2|) \quad (30)$$

$$f_y(x, y, u) < -c_3 \quad (31)$$

$$l_y(x, y, u) < c_4 \quad (32)$$

假设 2 存在常数 $c_5, c_6 > 0$, 使得

$$(y_1 - y_2)[f(x, y_1, u_1) - f(x, y_2, u_2)] \leq -c_5|y_1 - y_2|^2 + c_6|y_1 - y_2| \cdot |u_1 - u_2| \quad (33)$$

引理 1 设假设 1 和假设 2 成立, 令 $y, \tilde{y}$ 分别满足

$$Ay = h^2 F(y, u), A\tilde{y} = h^2 F(\tilde{y}, \tilde{u}) \quad (34)$$

则存在正常数 a_1 使得

$$\|y - \tilde{y}\|_w^2 \leq a_1 \|u - \tilde{u}\|_w^2 \quad (35)$$

证 将(34)中的两式相减可得

$$A(y - \tilde{y}) = h^2 F(y, u) - h^2 F(\tilde{y}, \tilde{u}) \quad (36)$$

上式与 $y - \tilde{y}$ 作内积, 并使用假设 2 可得

$$\begin{aligned} \langle y - \tilde{y}, A(y - \tilde{y}) \rangle &= h^2 \langle y - \tilde{y}, F(y, u) - F(\tilde{y}, \tilde{u}) \rangle \\ &\leq -hc_5 \|y - \tilde{y}\|_w^2 + hc_6 \|y - \tilde{y}\|_w \cdot \|u - \tilde{u}\|_w \end{aligned} \quad (37)$$

由于这里的矩阵 A 是对角占优矩阵, 且对角线元素都是正数, 所以对于任意非零向量 z , 有 $z^T Az \geq 0$, 故

$$0 \leq -hc_5 \|y - \tilde{y}\|_w^2 + hc_6 \|y - \tilde{y}\|_w \cdot \|u - \tilde{u}\|_w \quad (38)$$

$$\begin{aligned} &p_0[(1 + h\alpha)z_0 - z_1] + \sum_{n=1}^{N-1} p_n(-z_{n-1} + 2z_n - z_{n+1}) + p_N[-z_{N-1} + (1 - h\beta)z_N] \\ &= z_0[(1 + h\alpha)p_0 - p_1] + \sum_{n=1}^{N-1} z_n(-p_{n-1} + 2p_n - p_{n+1}) + z_N[-p_{N-1} + (1 - h\beta)p_N] \end{aligned} \quad (44)$$

证

$$\begin{aligned} &p_0[(1 + h\alpha)z_0 - z_1] + \sum_{n=1}^{N-1} p_n(-z_{n-1} + 2z_n - z_{n+1}) + p_N[-z_{N-1} + (1 - h\beta)z_N] \\ &= p_0[(1 + h\alpha)z_0 - z_1] + p_1(-z_0 + 2z_1 - z_2) + p_2(-z_1 + 2z_2 - z_3) + \cdots + \\ &\quad p_{N-1}(-z_{N-2} + 2z_{N-1} - z_N) + p_N[-z_{N-1} + (1 - h\beta)z_N] \\ &= z_0[(1 + h\alpha)p_0 - p_1] + z_1(-p_0 + 2p_1 - p_2) + z_2(-p_1 + 2p_2 - p_3) + \cdots + \\ &\quad z_{N-1}(-p_{N-2} + 2p_{N-1} - p_N) + z_N[-p_{N-1} + (1 - h\beta)p_N] \\ &= z_0[(1 + h\alpha)p_0 - p_1] + \sum_{n=1}^{N-1} z_n(-p_{n-1} + 2p_n - p_{n+1}) + z_N[-p_{N-1} + (1 - h\beta)p_N] \end{aligned} \quad (45)$$

定理 1 设 (y^k, u^k) 和 (y^{k+1}, u^{k+1}) 由 SQH 方法生成, u^k 和 u^{k+1} 是可测的, 令假设 1 成立. 则存在

所以

$$\|y - \tilde{y}\|_w \leq \frac{c_6}{c_5} \|u - \tilde{u}\|_w \quad (39)$$

引理 2 设假设 1 以及假设 2 成立, 对给定的 $u_n \in K_{ad}$, 设 p_n 由式(16) ~ (18) 生成, 则 p_n 为有界的.

证 将 p 与 $B(y, u)p$ 作内积并利用假设 1 可得

$$\begin{aligned} h \langle p, B(y, u)p \rangle &= h p^T B(y, u)p \\ &= h \sum_{n=0}^N w_n f_y(x_n, y_n, u_n) p_n^2 \leq -c_3 h \sum_{n=0}^N w_n p_n^2 \\ &= -c_3 \|p\|_w^2 \end{aligned} \quad (40)$$

将 p 与式(27)作内积可得

$$\begin{aligned} 0 \leq \langle p, Ap \rangle &= h^2 \langle p, B(y, u)p \rangle + h^2 \langle p, D(y, u) \rangle \\ &\leq -c_3 h \|p\|_w^2 + h \|p\|_w \cdot \|E(y, u)\|_w \end{aligned} \quad (41)$$

式中, $E(y, u) = [l_y(x_0, y_0, u_0), \dots, l_y(x_N, y_N, u_N)]^T$, 因此

$$\|p\|_w \leq \frac{1}{c_3} \|E(y, u)\|_w \quad (42)$$

将上式平方, 并利用假设 1 中的 $l_y(x, y, u) < c_4$, 可得

$$\begin{aligned} \|p\|_w^2 &\leq \frac{1}{c_3^2} \|E(y, u)\|_w^2 \\ &\leq \frac{h}{c_3^2} \sum_{n=0}^N w_n l_y^2(x_n, y_n, u_n) \\ &\leq \frac{hNc_4^2}{c_3^2} = \frac{(b-a)c_4^2}{c_3^2} \end{aligned} \quad (43)$$

所以 p 有界.

引理 3 以下等式恒成立

一个与 ϵ, k 和 u^k 无关的常数 $\theta > 0$, 对于算法 1 当前选择的 $\epsilon > 0$, 有

$$J_N(\mathbf{y}^{k+1}, \mathbf{u}^{k+1}) - J_N(\mathbf{y}^k, \mathbf{u}^k) \leq -\eta\tau. \quad (46)$$

证 由离散目标泛函的定义我们有

特别地, 对于 $\epsilon \geq \theta + \eta$ 和 $\tau = \|\mathbf{u}^{k+1} - \mathbf{u}^k\|_w^2$,

$$J_N(\mathbf{y}^{k+1}, \mathbf{u}^{k+1}) - J_N(\mathbf{y}^k, \mathbf{u}^k) = h \sum_{n=0}^N \omega_n [l(x_n, y_n^{k+1}, u_n^{k+1}) - l(x_n, y_n^k, u_n^k)] = h \sum_{n=0}^N \omega_n [H(x_n, y_n^{k+1}, u_n^{k+1}, p_n^k) - H(x_n, y_n^k, u_n^k, p_n^k)] - h \sum_{n=0}^N \omega_n [p_n^k f(x_n, y_n^{k+1}, u_n^{k+1}) - p_n^k f(x_n, y_n^k, u_n^k)] \quad (47)$$

易知

$$J_N(\mathbf{y}^{k+1}, \mathbf{u}^{k+1}) - J_N(\mathbf{y}^k, \mathbf{u}^k) = h \sum_{n=0}^N \omega_n [H(x_n, y_n^{k+1}, u_n^{k+1}, p_n^k) - H(x_n, y_n^k, u_n^{k+1}, p_n^k)] + h \sum_{n=0}^N \omega_n [H(x_n, y_n^k, u_n^k, p_n^k) - H(x_n, y_n^k, u_n^k, p_n^k)] - h \sum_{n=0}^N \omega_n [p_n^k f(x_n, y_n^{k+1}, u_n^{k+1}) - p_n^k f(x_n, y_n^k, u_n^k)] \quad (48)$$

对于式(48)中的第一项 $H(x_n, y_n^{k+1}, u_n^{k+1}, p_n^k) - H(x_n, y_n^k, u_n^{k+1}, p_n^k)$, 我们有

$$\begin{aligned} H(x_n, y_n^{k+1}, u_n^{k+1}, p_n^k) - H(x_n, y_n^k, u_n^{k+1}, p_n^k) &= \int_0^1 H_y[x_n, y_n^k + s(y_n^{k+1} - y_n^k), u_n^{k+1}, p_n^k](y_n^{k+1} - y_n^k) ds \\ &= \int_0^1 \{H_y[x_n, y_n^k + s(y_n^{k+1} - y_n^k), u_n^{k+1}, p_n^k] - H_y(x_n, y_n^k, u_n^k, p_n^k)\}(y_n^{k+1} - y_n^k) ds + H_y(x_n, y_n^k, u_n^k, p_n^k)(y_n^{k+1} - y_n^k) \\ &= \int_0^1 \{f_y[x_n, y_n^k + s(y_n^{k+1} - y_n^k), u_n^{k+1}] - f_y(x_n, y_n^k, u_n^k)\} p_n^k (y_n^{k+1} - y_n^k) ds + \{l_y[x_n, y_n^k + s(y_n^{k+1} - y_n^k), u_n^{k+1}] - l_y(x_n, y_n^k, u_n^k)\}(y_n^{k+1} - y_n^k) ds + H_y(x_n, y_n^k, u_n^k, p_n^k)(y_n^{k+1} - y_n^k) \\ &\leq \int_0^1 |f_y[x_n, y_n^k + s(y_n^{k+1} - y_n^k), u_n^{k+1}] - f_y(x_n, y_n^k, u_n^k)| \cdot |p_n^k| \cdot |y_n^{k+1} - y_n^k| ds + \int_0^1 |l_y[x_n, y_n^k + s(y_n^{k+1} - y_n^k), u_n^{k+1}] - l_y(x_n, y_n^k, u_n^k)| \cdot |y_n^{k+1} - y_n^k| ds + H_y(x_n, y_n^k, u_n^k, p_n^k)(y_n^{k+1} - y_n^k) \end{aligned} \quad (49)$$

这里第三个等号中我们使用了 $H_y = f_y p + l_y$.

由假设 1 中的 Lipschitz 条件, 我们有:

$$\begin{aligned} H(x_n, y_n^{k+1}, u_n^{k+1}, p_n^k) - H(x_n, y_n^k, u_n^{k+1}, p_n^k) &\leq \int_0^1 c_1 [|s(y_n^{k+1} - y_n^k)| + |u_n^{k+1} - u_n^k|] |p_n^k| \cdot |y_n^{k+1} - y_n^k| ds + \int_0^1 c_2 [|s(y_n^{k+1} - y_n^k)| + |u_n^{k+1} - u_n^k|] |y_n^{k+1} - y_n^k| ds + H_y(y_n^k, u_n^k, p_n^k)(y_n^{k+1} - y_n^k) \\ &= |p_n^k| \cdot c_1 [\frac{1}{2}(y_n^{k+1} - y_n^k)^2 + |u_n^{k+1} - u_n^k| \cdot |y_n^{k+1} - y_n^k|] + c_2 [\frac{1}{2}(y_n^{k+1} - y_n^k)^2 + |u_n^{k+1} - u_n^k| \cdot |y_n^{k+1} - y_n^k|] + H_y(x_n, y_n^k, u_n^k, p_n^k)(y_n^{k+1} - y_n^k) \\ &\leq \frac{c_7 c_1 + c_2}{2} (y_n^{k+1} - y_n^k)^2 + (c_7 c_1 + c_2) |u_n^{k+1} - u_n^k| \cdot |y_n^{k+1} - y_n^k| + H_y(x_n, y_n^k, u_n^k, p_n^k)(y_n^{k+1} - y_n^k) \end{aligned} \quad (50)$$

式(50)最后一个不等式利用了引理 2, 这里我们不妨设 $|p_n| \leq c_7$. 同时我们有

$$\begin{aligned} h \sum_{n=0}^N \omega_n [H(x_n, y_n^{k+1}, u_n^{k+1}, p_n^k) - H(x_n, y_n^k, u_n^{k+1}, p_n^k)] &\leq \frac{c_7 c_1 + c_2}{2} h \sum_{n=0}^N \omega_n (y_n^{k+1} - y_n^k)^2 + (c_7 c_1 + c_2) h \sum_{n=0}^N \omega_n |u_n^{k+1} - u_n^k| \cdot |y_n^{k+1} - y_n^k| + h \sum_{n=0}^N \omega_n H_y(x_n, y_n^k, u_n^k, p_n^k)(y_n^{k+1} - y_n^k) \\ &\leq \frac{c_6 (c_7 c_1 + c_2)}{2 c_5} h \sum_{n=0}^N \omega_n (u_n^{k+1} - u_n^k)^2 + \frac{2 c_6 (c_7 c_1 + c_2)}{c_5} h \sum_{n=0}^N \omega_n (u_n^{k+1} - u_n^k)^2 + h \sum_{n=0}^N \omega_n H_y(x_n, y_n^k, u_n^k, p_n^k)(y_n^{k+1} - y_n^k) \\ &= \frac{5 c_6 (c_7 c_1 + c_2)}{2 c_5} h \sum_{n=0}^N \omega_n (u_n^{k+1} - u_n^k)^2 + h \sum_{n=0}^N \omega_n H_y(x_n, y_n^k, u_n^k, p_n^k)(y_n^{k+1} - y_n^k) \\ &= \frac{5 c_6 (c_7 c_1 + c_2)}{2 c_5} \|\mathbf{u}^{k+1} - \mathbf{u}^k\|_w^2 + h \sum_{n=0}^N \omega_n H_y(x_n, y_n^k, u_n^k, p_n^k)(y_n^{k+1} - y_n^k) \end{aligned} \quad (51)$$

式(51)最后一个不等式利用了引理 1 和柯西-施瓦茨不等式.

接下来,我们计算 $h \sum_{n=0}^N \omega_n H_y(x_n, y_n^k, u_n^k, p_n^k)(y_n^{k+1} - y_n^k)$ 的值

$$\begin{aligned} h \sum_{n=0}^N \omega_n H_y(x_n, y_n^k, u_n^k, p_n^k)(y_n^{k+1} - y_n^k) &= h \sum_{n=0}^N \omega_n [f_y(x_n, y_n^k, u_n^k) p_n^k + l_y(x_n, y_n^k, u_n^k)] \varphi_n \\ &= \frac{1}{h} \left\{ [(1+h\alpha)p_0^k - p_1^k] \varphi_0 + \sum_{n=1}^{N-1} (-p_{n-1}^k + 2p_n^k - p_{n+1}^k) \varphi_n + [-p_{N-1}^k + (1-h\beta)p_N^k] \varphi_N \right\} \\ &= \frac{1}{h} \left\{ p_0^k [(1+h\alpha)\varphi_0 - \varphi_1] + \sum_{n=1}^{N-1} p_n^k (-\varphi_{n-1} + 2\varphi_n - \varphi_{n+1}) + p_N^k [-\varphi_{N-1} + (1-h\beta)\varphi_N] \right\} \end{aligned} \quad (52)$$

其中 $\varphi_n = y_n^{k+1} - y_n^k, n=0, 1, \dots, N$. 式(52)第二个等式由式(29)~(31)得出且第三个等式由引理 3 得出.

由离散 SQH 算法,我们有

$$H_\epsilon(x_n, y_n^k, u_n^{k+1}, u_n^k, p_n^k) = \min_{u \in K_{ad}} H_\epsilon(x_n, y_n^k, u, u_n^k, p_n^k) \quad (53)$$

因此由式(48)的第二项,可得

$$\begin{aligned} h \sum_{n=0}^N \omega_n [H(x_n, y_n^k, u_n^{k+1}, p_n^k) - H(x_n, y_n^k, u_n^k, p_n^k)] \\ = h \sum_{n=0}^N \omega_n [H_\epsilon(x_n, y_n^k, u_n^{k+1}, u_n^k, p_n^k) - \epsilon(u_n^{k+1} - u_n^k)^2 - H_\epsilon(x_n, y_n^k, u_n^k, u_n^k, p_n^k) + \epsilon(u_n^k - u_n^k)^2] \\ \leq h \sum_{n=0}^N \omega_n [-\epsilon(u_n^{k+1} - u_n^k)^2] = -\epsilon \|u^{k+1} - u^k\|_w^2 \end{aligned} \quad (54)$$

由式(48)的第三项,可得

$$\begin{aligned} -h \sum_{n=0}^N \omega_n [p_n^k f(x_n, y_n^{k+1}, u_n^{k+1}) - p_n^k f(x_n, y_n^k, u_n^k)] &= -\frac{1}{h} p_0^k \{[(1+h\alpha)y_0^{k+1} - y_1^{k+1} + hg_1] - \\ &[(1+h\alpha)y_0^k - y_1^k + hg_1]\} - \frac{1}{h} \sum_{n=1}^{N-1} p_n^k [(-y_{n-1}^{k+1} + 2y_n^{k+1} - y_{n+1}^{k+1}) - (-y_{n-1}^k + 2y_n^k - y_{n+1}^k)] - \\ &\frac{1}{h} p_N^k \{[-y_{N-1}^{k+1} + (1-h\beta)y_N^{k+1} - hg_2] - [-y_{N-1}^k + (1-h\beta)y_N^k - hg_2]\} \\ &= -\frac{1}{h} \{p_0^k [(1+h\alpha)\varphi_0 - \varphi_1] + \sum_{n=1}^{N-1} p_n^k (-\varphi_{n-1} + 2\varphi_n - \varphi_{n+1}) + p_N^k [-\varphi_{N-1} + (1-h\beta)\varphi_N]\} \end{aligned} \quad (55)$$

所以

$$\begin{aligned} J_N(y^{k+1}, u^{k+1}) - J_N(y^k, u^k) &\leq \frac{5c_6(c_7c_1 + c_2)}{2c_5} \|u^{k+1} - u^k\|_w^2 + \frac{1}{h} \{p_0^k [(1+h\alpha)\varphi_0 - \varphi_1] + \\ &\sum_{n=1}^{N-1} p_n^k (-\varphi_{n-1} + 2\varphi_n - \varphi_{n+1}) + p_N^k [-\varphi_{N-1} + (1-h\beta)\varphi_N]\} + (-\epsilon) \|u^{k+1} - u^k\|_w^2 - \\ &\frac{1}{h} \{p_0^k [(1+h\alpha)\varphi_0 - \varphi_1] + \sum_{n=1}^{N-1} p_n^k (-\varphi_{n-1} + 2\varphi_n - \varphi_{n+1}) + p_N^k [-\varphi_{N-1} + (1-h\beta)\varphi_N]\} \\ &= \frac{5c_6(c_7c_1 + c_2)}{2c_5} \|u^{k+1} - u^k\|_w^2 - \epsilon \|u^{k+1} - u^k\|_w^2 = -(\epsilon - \theta) \|u^{k+1} - u^k\|_w^2 \end{aligned} \quad (56)$$

式中, $\theta = \frac{5c_6(c_7c_1 + c_2)}{2c_5}$. 从而

$$\begin{aligned} J_N(y^{k+1}, u^{k+1}) - J_N(y^k, u^k) \\ \leq -(\epsilon - \theta) \|u^{k+1} - u^k\|_w^2 \end{aligned} \quad (57)$$

定理 2 设定理 1 的假设成立,如果在离散 SQH 算法中的每次迭代时选择 $\epsilon = \theta + \eta$,则

(a) 序列 $\{J_N(y^k, u^k)\}_{k=0,1,2,\dots}$ 是单调递减的,

并且收敛.

(b) $\lim_{k \rightarrow \infty} \|u^{k+1} - u^k\|_w^2 = 0$.

证 证明(a)根据定理 1,可得 $J_N(y^{k+1}, u^{k+1}) \leq J_N(y^k, u^k)$,所以序列 $\{J_N(y^k, u^k)\}_{k=0,1,2,\dots}$ 是单调递减的.再由算法 1 第六步可知收敛.

为了证明(b),我们将式(46)重写为

$$\begin{aligned} \|u^{k+1} - u^k\|_w^2 \\ \leq \frac{1}{\eta} [J_N(y^k, u^k) - J_N(y^{k+1}, u^{k+1})] \end{aligned} \quad (58)$$

上式关于 $k=0, 1, \dots, K$ 求和可得

$$\sum_{k=0}^K \| \mathbf{u}^{k+1} - \mathbf{u}^k \|_w^2 \leq \frac{1}{\eta} [J_N(\mathbf{y}^0, \mathbf{u}^0) - J_N(\mathbf{y}^{K+1}, \mathbf{u}^{K+1})] \tag{59}$$

这表明当 $K \rightarrow \infty$ 时, 具有正元素的级数

$\sum_{k=0}^K \| \mathbf{u}^{k+1} - \mathbf{u}^k \|_w^2$ 是收敛的, 因此证明了(b).

4 数值实验

在本节中, 我们从数值上研究离散 SQH 算法的收敛性. 我们采用 MATLAB 软件进行数值计算. 离散 SQH 方法中的参数均设置为 $\kappa=10^{-8}, \sigma=1.2, \zeta=0.9, \eta=10^{-5}, \epsilon=1$ 且初始控制 $\mathbf{u}^0$ 取 0.

考虑如下椭圆最优控制问题

$$\begin{aligned} \min J(y, u) = \min \int_0^1 & \left[\frac{1}{2} y^2(x) + \frac{m_1}{2} u^2(x) + m_2 u(x) \right] dx \\ \text{s. t. } & \begin{cases} -y'' = -y + u, & u \in K_{\text{ad}} \\ y'(0) = g_1 \\ y'(1) = g_2 \end{cases} \end{aligned} \tag{60}$$

这里我们取 $m_1=10^{-3}, m_2=10^{-6}, g_1=g_2=1, K_{\text{ad}}=[-1, 1], h=1/800$.

在图 1 中, 我们展示了离散 SQH 方法得到的最优控制 u 及相应状态变量和控制变量随 x 的变化曲线. 在图 2 中, 我们展示了离散目标泛函 J_N 随 SQH 迭代的演化, 可以看到离散目标泛函是单调递减且收敛的, 这验证了我们理论结果的正确性.

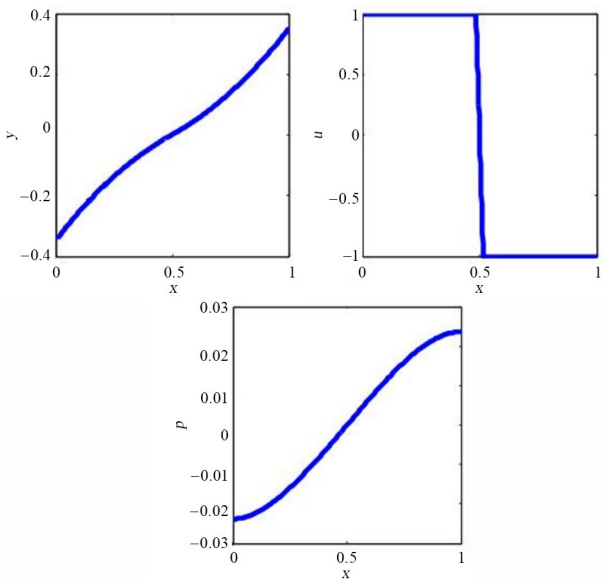


图 1 状态变量 y 、控制变量 u 和伴随变量 p 关于自变量 x 的变化曲线
Fig. 1 The curves of the state variable y , control variable u , and adjoint variable p with respect to the independent variable x

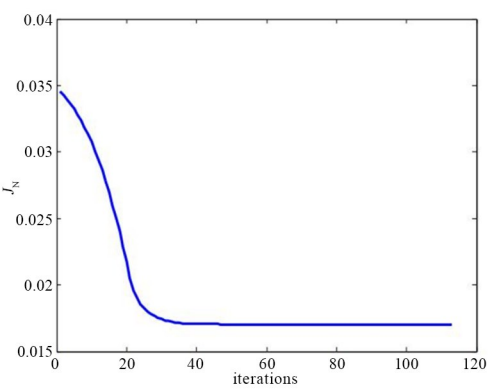


图 2 离散目标函数 J_N 随 SQH 迭代的演化
Fig. 2 The evolution of J_N along the SQH iterations

为了进一步验证该算法的有效性, 我们接下来考虑如下具有非光滑目标泛函和非线性微分方程约束的最优控制问题

$$\begin{aligned} \min J(y, u) = \min \int_0^1 & \left\{ \frac{1}{2} [y(x) - y_d]^2 + \frac{m_3}{2} u^2(x) + m_4 |u(x)| \right\} dx \\ \text{s. t. } & \begin{cases} -y'' = -y^3 - y + u, & u \in K_{\text{ad}} \\ y'(0) = g_3 \\ y'(1) = g_4 \end{cases} \end{aligned} \tag{61}$$

这里我们取 $m_3=10^{-8}, m_4=10^{-6}, g_3=g_4=0, y_d=0.8\sin(\pi x), K_{\text{ad}}=[-1, 1], h=1/100$.

在图 3 中, 我们展示了离散 SQH 方法得到的最优控制 u 及相应状态变量和控制变量随 x 的变化曲线. 在图 4 中, 我们展示了离散目标泛函 J_N 随 SQH 迭代的演化, 可以看到离散目标泛函是单调递减且收敛的, 这验证了我们理论结果的正确性.

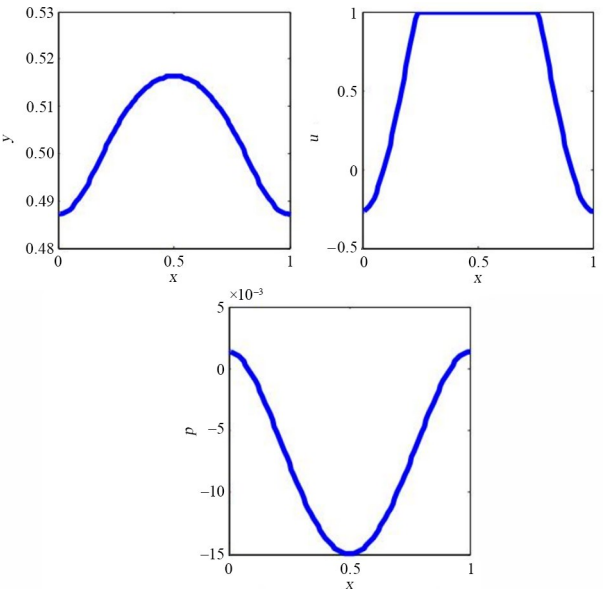


图 3 状态变量 y 、控制变量 u 和伴随变量 p 关于自变量 x 的变化曲线
Fig. 3 The curves of the state variable y , control variable u , and adjoint variable p with respect to the independent variable x

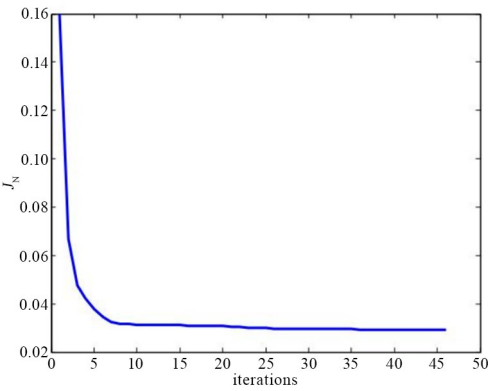


图 4 离散目标函数 J_N 随 SQH 迭代的演化
Fig. 4 The evolution of J_N along the SQH iterations

为了验证离散 SQH 算法的效率,我们将其与 Adam 优化算法^[26]、自适应梯度下降法^[27] (adaptive gradient descent, AdaGD)进行了数值比较. 这里我们考虑最优控制问题(60). 图 5 展示了三种优化算法在离散目标函数值 J_N 收敛过程中的表现. 可以发现,离散 SQH 方法比 Adam 方法收敛速度更快, SQH 方法与 Adam 方法收敛速度相当.

表 2 对比了 SQH 算法、Adam 算法以及 AdaGD 算法在目标函数达到不同数值时 CPU 计算时间和迭代次数,其中,Iter 表示迭代次数. 实验结果表明,

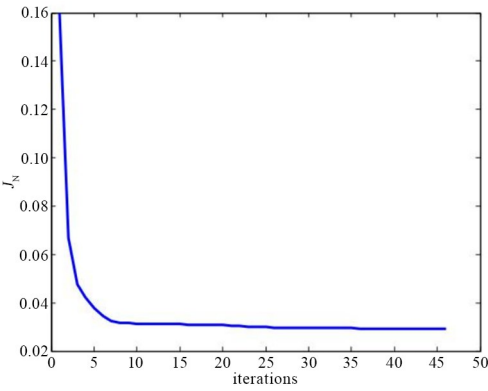


图 4 离散目标函数 J_N 随 SQH 迭代的演化
Fig. 4 The evolution of J_N along the SQH iterations

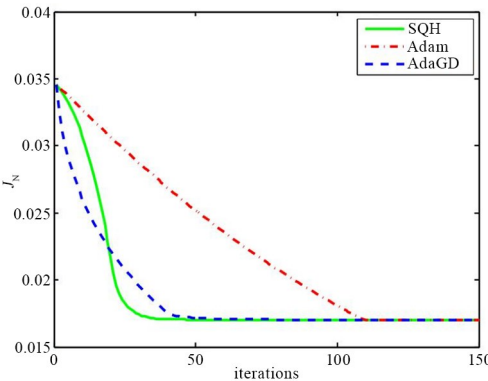


图 5 离散目标泛函 J_N 随迭代的演化
Fig. 5 The evolution of J_N along the iterations

表 2 SQH 算法、Adam 算法与 AdaGD 算法比较 (CPU 单位:秒)						
Table 2 Comparison of SQH, Adam, and AdaGD algorithms (CPU Unit: seconds)						
J_N	SQH CPU	Iter	Adam CPU	Iter	AdaGD CPU	Iter
0.018	0.0057	25	0.0257	101	0.0085	38
0.020	0.0048	21	0.0195	85	0.0068	29
0.023	0.0041	18	0.0151	64	0.0042	18
0.026	0.0037	16	0.0107	46	0.0026	11
0.029	0.0028	12	0.0069	29	0.0015	6
0.032	0.0017	7	0.0034	14	0.0008	3
0.034	0.0007	2	0.0011	4	0.0005	2

在目标函数达到相同数值时 SQH 算法的 CPU 时间相比 Adam 和 AdaGD 方法更少.

5 总结

本文针对一维椭圆最优控制问题,在“先离散再优化”框架下,构造了一类基于中心差分格式的离散序列二次哈密顿方法,并证明了该离散格式能保持连续情形下序列二次哈密顿迭代的单调收敛性. 数值实验验证了理论分析的正确性. 在后续研究中,我们将考虑高维椭圆最优控制问题的离散序列二次哈密顿方法.

参考文献

[1] ABDULLA U G, BUKSHTYNOV V, SEIF S. Cancer detection through Electrical Impedance Tomography and optimal control theory: theoretical and computational analysis [J]. Mathematical Biosciences and Engineering, 2021, 18(4): 4834 — 4859.

[2] GRASS D, UECKER H, UPMANN T. Optimal fishery with coastal catch [J]. Natural Resource Modeling, 2019, 32(4): e12235.

[3] CIAMBELLA J, TOMASSETTI G. A form-finding strategy for magneto-elastic actuators [J]. International Journal of Non-Linear Mechanics, 2020, 119: 103297.

[4] PONTRYAGIN L S, BOLTYANSKI V G, GAMKRELIDZE R V, et al. The mathematical theory of optimal processes [M]. New York: John Wiley & Sons, 1962.

[5] BELLMAN R. Dynamic programming[M]. Prince-

- ton; Princeton University Press, 1957.
- [6] KALMAN R E. On the general theory of control systems [C]//Proceedings of the First International Congress of the International Federation of Automatic Control (Moscow, 1960). London: Butterworth, 1961: 481—492.
 - [7] GERDTS M. Direct shooting method for the numerical solution of higher-index DAE optimal control problems [J]. *Journal of Optimization Theory and Applications*, 2003, 117(2): 267—294.
 - [8] BERGOUNIOUX M, KUNISCH K. Primal-dual strategy for state-constrained optimal control problems [J]. *Computational Optimization and Applications*, 2002, 22(2): 193—224.
 - [9] ULBRICH M. Semismooth Newton methods for operator equations in function spaces [J]. *SIAM Journal on Optimization*, 2002, 13(3): 805—841.
 - [10] BÜSKENS C, MAURER H. SQP-methods for solving optimal control problems with control and state constraints; adjoint variables, sensitivity analysis and real-time control [J]. *Journal of Computational and Applied Mathematics*, 2000, 120(1/2): 85—108.
 - [11] TRÖLTZSCH F. On the Lagrange-Newton-SQP method for the optimal control of semilinear parabolic equations [J]. *SIAM Journal on Control and Optimization*, 1999, 38(1): 294—312.
 - [12] BERGOUNIOUX M. Augmented Lagrangian method for distributed optimal control problems with state constraints [J]. *Journal of Optimization Theory and Applications*, 1993, 78(3): 493—521.
 - [13] BOYD S, PARIKH N, CHU E, et al. Distributed optimization and statistical learning via the alternating direction method of multipliers[J]. *Now Publishers*, 2011, 3(1): 1—122.
 - [14] FALK R S. Approximation of a class of optimal control problems with order of convergence estimates [J]. *Journal of Mathematical Analysis and Applications*, 1973, 44(1): 28—47.
 - [15] CHEN Y P, YI N Y, LIU W B. A Legendre-Galerkin spectral method for optimal control problems governed by elliptic equations [J]. *SIAM Journal on Numerical Analysis*, 2008, 46(5): 2254—2275.
 - [16] LISSY P, ROVENTA I. Optimal approximation of internal controls for a wave-type problem with fractional Laplacian using finite-difference method [J]. *Mathematical Models and Methods in Applied Sciences*, 2020, 30(3): 439—475.
 - [17] CHEN Y P, DAI Y Q. Superconvergence for optimal control problems governed by semi-linear elliptic equations [J]. *Journal of Scientific Computing*, 2009, 39(2): 206—221.
 - [18] HUANG F L, CHEN Y P. Error estimates for spectral approximation of elliptic control problems with integral state and control constraints [J]. *Computers & Mathematics with Applications*, 2014, 68(8): 789—803.
 - [19] LIN X X, CHEN Y P, HUANG Y Q. Galerkin spectral approximation of optimal control problems with L^2 -norm control constraint [J]. *Applied Numerical Mathematics*, 2020, 150: 418—432.
 - [20] LI B J, LIU S Y. Conjugate gradient-boundary element solution for distributed elliptic optimal control problems [J]. *Journal of Mathematical Analysis and Applications*, 2007, 335(2): 1219—1237.
 - [21] SONG X L, YU B, WANG Y Y, et al. An FE-inexact heterogeneous ADMM for elliptic optimal control problems with L^1 -control cost [J]. *Journal of Systems Science and Complexity*, 2018, 31(6): 1659—1697.
 - [22] ZHANG Z M, CHEN X S. A conjugate gradient method for distributed optimal control problems with nonhomogeneous Helmholtz equation [J]. *Applied Mathematics and Computation*, 2021, 402: 126019.
 - [23] BORZÌ A. The sequential quadratic Hamiltonian method; solving optimal control problem [M]. Boca Raton, FL, USA: CRC Press, 2023, 45—185.
 - [24] BREITENBACH T, BORZÌ A. On the SQH scheme to solve nonsmooth PDE optimal control problems [J]. *Numerical Functional Analysis and Optimization*, 2019, 40(13): 1489—1531.
 - [25] BONNANS J F. On an algorithm for optimal control using Pontryagin's maximum principle [J]. *SIAM Journal on Control and Optimization*, 1986, 24(3): 579—588.
 - [26] KINGMA D P, BA J, HAMMAD M M. Adam: a method for stochastic optimization [EB/OL]. <https://arxiv.org/abs/1412.6980v9>.
 - [27] DUCHI J, AZAN E, INGER Y. Adaptive sub-gradient methods for online learning and stochastic optimization [J]. *Journal of Machine Learning Research*, 2011, 12: 2121—2159.