

带罗宾边界椭圆最优控制问题的自适应交替乘子方向方法求解

林星源

广东工业大学数学与统计学院, 广东 广州

收稿日期: 2026年3月13日; 录用日期: 2026年4月10日; 发布日期: 2026年4月22日

摘要

本文研究带罗宾边界条件的椭圆型最优控制问题。基于有限元离散化方法, 将该问题转化为优化问题, 并采用自适应乘子交替方向法(adaptive-ADMM)对其进行数值求解。在不依赖拉格朗日乘子存在性的前提下, 证明了交替乘子方向方法的收敛性。此外, 利用问题对惩罚参数 ρ 的敏感性变化, 提出了 ρ 的自适应调整策略, 显著提升了ADMM方法的鲁棒性与收敛速度。本文同时建立了遍历型与非遍历型收敛速率估计, 并将所提自适应交替乘子方向方法与采用不同固定值的标准交替乘子方向方法进行对比, 验证了该方法的有效性。

关键词

最优控制, 罗宾边界, 自适应交替乘子方向法, 收敛性分析

An Adaptive for Elliptic Optimal Control Problems with Robin Boundary Conditions

Xingyuan Lin

School of Mathematics and Statistics, Guangdong University of Technology, Guangzhou
Guangdong

Abstract

This paper investigates elliptic optimal control problems with Robin boundary conditions. Based on the finite element discretization method, the problem is transformed into an optimization problem and solved numerically using the adaptive alternating direction method of multipliers (adaptive-ADMM). Without relying on the existence of Lagrange multipliers, the convergence of the ADMM method is proved. Furthermore, by exploiting the sensitivity of the problem to the penalty parameter ρ , an adaptive adjustment strategy for ρ is proposed, which significantly improves the robustness and convergence speed of the ADMM method. Both ergodic and non-ergodic convergence rate estimates are established. The proposed adaptive ADMM method is compared with the standard ADMM method using different fixed parameter values, and the numerical results verify its effectiveness.

Keywords

Optimal Control, Robin Boundary, Adaptive Alternating Direction Method of Multipliers, Convergence Analysis

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 前言

近年来, 含偏微分方程的最优控制问题在工程、生物医学和经济学领域得到了广泛应用。Mehdaoui [1]研究了带状态约束的耦合拟线性抛物型系统的最优控制问题, 分析了最优解的存在性并推导了一阶必要最优性条件, 所建模型用于刻画技术进步下的经济增长问题。Enaet Hossain [2]基于分裂改进迎风格式和差分进化算法, 建立了多孔介质中多组分污染的偏微分方程约束最优控制模型, 通过控制污染物注入速率最小化模拟浓度与观测浓度的误差, 并利用数值实验验证了模型的有效性。Li等人[3]设计了多列高速列车运行的分布式最优控制方法, 该模型克服了通信约束, 实现了列车速度的有效调控。

乘子交替方向法(ADMM)固有的解耦策略使其具备简洁、高效的优点,因此被广泛应用于各类优化问题与最优控制问题中。学者们针对具体问题提出了多种ADMM改进算法。Pierer von Esch等人[4]提出一种改进ADMM算法,通过交换协方差信息处理具有非线性动力学特性的随机多智能体系统,将随机优化问题转化为确定性形式,实现了问题的分布式求解。Chen [5]提出一种多层乘子交替方向法,采用“优化-离散-再优化”的求解策略,先在函数空间中应用交替乘子法,再通过网格逐步加密的有限元方法对子问题进行离散化,同时融入误差控制以保证算法的收敛性。Li等人[6]提出基于舒尔补的半邻近乘子交替方向法,实现了凸二次锥规划及其扩展问题的高效求解,该方法在处理多块优化结构时能提供收敛性保证,并通过数值实验验证了其在二次半定规划等问题上的优越性能。与多块优化相关研究类似,Chen [7]证明,对于包含三个及以上分块的可分离凸优化问题,乘子交替方向法的直接扩展形式并非必然收敛,他们构造了具体的反例验证该扩展形式的发散性,并给出了保证其收敛的充分条件。Song等人[8]针对非光滑偏微分方程约束优化问题,提出ADMM-PINNs算法框架,通过交替乘子法对问题进行解耦,使物理信息神经网络(PINNs)能够处理非光滑项,拓宽了该网络的适用范围。

如今,自适应交替乘子方向法的策略已经多种多样。Song和Wang等人[9]提出A-ADMM,通过动态调整全局惩罚参数,依据原始残差与对偶残差的比例加速收敛,适用于多主体协同优化的分布式场景。Chen和He等人[10]提出FedADMM-InSa,不仅设计了基于梯度残差的不精确性准则以降低客户端计算负载,还允许每个客户端独立自适应调整自身惩罚参数,以应对数据与系统异构性。Chen和Ye等人[11]提出的自适应共识交替方向乘子法(AC-ADMM)中,核心是动态调整惩罚参数,以提升分布式优化在异构多能源微电网(MEMG)协同中的收敛性能。本文使用的策略是通过变分分析和有限元离散将问题转化为向量-矩阵相关的优化问题再基于原始残差与对偶残差动态调整惩罚参数在该类最优控制问题中实现自适应ADMM。本文自适应策略在应用场景上具有优势,属于偏微分方程约束优化领域,而非分布式优化或信号处理,并且将问题从简单边界扩展到复杂边界上。

边界条件对偏微分方程最优控制问题数值解的精度以及最优解的存在性和唯一性均具有重要影响。在最优控制领域,常用的边界条件包括狄利克雷(Dirichlet)、诺伊曼(Neumann)和罗宾(Robin)边界条件(前两种边界条件的组合形式)。狄利克雷边界条件作为最经典的边界条件之一,已被众多学者展开深入研究。文献[12]针对分布式椭圆型最优控制问题提出了一种自适应有限元方法,该方法融入变能量正则化技术,通过局部误差指标驱动网格的自适应加密,数值实验验证了该方法在二维和三维问题中均具有良好的效率与鲁棒性。Banz [13]采用hp有限元对带控制约束的椭圆型最优控制问题进行离散,推导了先验和后验误差估计,并通过hp自适应网格加密显著提升了收敛效率。Brenner [14]将问题重新转化为四阶变分不等式并推导了先验误差估计,同时提出了基于残量的后验误差估计子和自适应网格加密算法。Yang和Sun[15]针对带诺伊曼边界条件的抛物型最优控制问题,利用二阶后向差分公式(BDF2)构建了二阶时间离散格式,推导了先验误差估计,并通过数值实验验证了该方法的二阶时间精度。Wang和Zhou[16]针对带罗宾边界条件的半线性椭圆型最优控制问题提出了多层校正有限元方法,该方法在保持最优收敛精度的同时,大幅降低了计算复杂度。Chrysafinos等人[17]研究了带罗宾边界条件的最优边界控制问题,构建了半离散有限元近似格式,并推导了最优性系统与误差估计。

本文采用乘子交替方向法研究含罗宾边界条件和边界观测的最优控制问题,丰富了该领域中

关于复杂边界条件与边界观测问题的研究内容。本文采用“先离散后优化”的求解策略，通过有限元离散将控制问题转化为优化问题，并选用交替乘子法作为该优化问题的求解方法。求解过程中，惩罚参数 ρ 的取值过大或过小均会导致算法收敛速度变慢，因此本文对 ρ 的选取方式进行改进，提出了自适应交替乘子方向法。开展收敛性分析，并通过两种不同的方法分别证明了其非遍历型与遍历型收敛速率。

本文的结构安排如下：第1节介绍研究背景并提出问题。第2节给出问题的有限元离散形式以及原始交替乘子方向法算法。第3节提供若干预备结果，证明交替乘子方向法的收敛性质以及遍历与非遍历收敛速率，同时提出改进的交替乘子方向法方法。第4节通过数值实验，将所提出的自适应交替乘子方向法方法与传统交替乘子方向法方法、一类带线搜索的共轭梯度法进行对比，验证了该方法的合理性与有效性。

椭圆型最优控制问题表述如下：

$$\min_{u \in U} J(y, u) = \frac{\lambda_\Omega}{2} \|y - y_\Omega\|_{L^2(\Omega)}^2 + \frac{\lambda_\Gamma}{2} \|y - y_\Gamma\|_{L^2(\Gamma)}^2 + \frac{\lambda}{2} \|u\|_{L^2(\Omega)}^2 \quad (1.1)$$

满足如下椭圆型偏微分方程：

$$-\Delta y = \beta u \text{ in } \Omega \quad (1.2)$$

$$\partial_n y + \alpha y = 0 \text{ on } \Gamma \quad (1.3)$$

以及控制约束

$$U_{ad} = \{u \in L^2(\Omega) \cap C^\omega(\Omega) | u_a \leq u \leq u_b, u_a, u_b \in \Omega\} \in U = L^2(\Omega).$$

设 Γ 是区域 Ω 的边界， n 表示 Ω 的外法向量， λ_Ω 和 λ_Γ 为非负常数， λ 严格正数。此外， $\beta \in L^\infty(\Omega)$ ， $\alpha \in L^\infty(\Gamma)$ 。其中， $\alpha > 0$ 在 Γ 上处处成立， M 为正数。目标函数 y_Γ 和 y_Ω 分别属于 $L^2(\Omega)$ ， $L^2(\Gamma)$ 。全文中，定理和引理中出现的通用常数均用 C 表示。

2. 应用交替方向乘子法求解(1.1)-(1.3)

在本章中，我们将通过变分学和有限元离散化对问题(1.1)-(1.3)进行变换，利用交替乘子方向法方法解耦问题以获得相应的子问题，并给出算法设计。为了研究离散化后的问题，区域 $\Omega \subset \mathbb{R}^2$ 被定义为一个凸区域。我们将控制空间设为 $U = L^2(\Omega)$ ，状态空间设为 $Y = H^1(\Omega)$ 。

引入向量化变量 $y \in \mathbb{R}^N$ 和 $u \in \mathbb{R}^N$ ，并引入体积质量矩阵 M_v 和边界目标质量矩阵 M_g （大小为 $N \times N$ 的对称矩阵），那么问题(1.1)变为：

$$\min J(\mathbf{y}, \mathbf{u}) = \frac{\lambda_\Omega}{2} (\mathbf{y} - \mathbf{y}_\Omega)^T M_v (\mathbf{y} - \mathbf{y}_\Omega) + \frac{\lambda_\Gamma}{2} (\mathbf{y} - \mathbf{y}_\Gamma)^T M_g (\mathbf{y} - \mathbf{y}_\Gamma) + \frac{\lambda}{2} \mathbf{u}^T M_v \mathbf{u},$$

利用格林公式对(1.2)-(1.3)进行处理，可得

$$\int_\Omega \nabla y \nabla v + \int_\Gamma \alpha y v = \int_\Omega \beta u v,$$

$$\int_{\Omega} \nabla y \nabla v + \int_{\Gamma} \alpha y v = \int_{\Omega} \beta u v,$$

引入边界质量矩阵 M_g (一个大小为 $N \times N$ 的对称矩阵), 经有限元离散化后可得

$$K\mathbf{y} + \alpha M\mathbf{y} = \beta M_v \mathbf{u}.$$

问题(1.1)-(1.3)可改写为

$$\min J(\mathbf{y}, \mathbf{u}) = \frac{\lambda_{\Omega}}{2} (\mathbf{y} - \mathbf{y}_{\Omega})^T M_v (\mathbf{y} - \mathbf{y}_{\Omega}) + \frac{\lambda_{\Gamma}}{2} (\mathbf{y} - \mathbf{y}_{\Gamma})^T M_g (\mathbf{y} - \mathbf{y}_{\Gamma}) + \frac{\lambda}{2} \mathbf{u}^T M_v \mathbf{u}, \quad (2.1)$$

$$K\mathbf{y} + \alpha M\mathbf{y} = \beta M_v \mathbf{u}, \quad (2.2)$$

根据问题(2.1)-(2.2), 我们可以得到对应的拉格朗日泛函: $\mathcal{L}_{\rho}: (U \times Y) \times Y \rightarrow \mathbb{R} \cup \{+\infty\}$

$$\begin{aligned} \mathcal{L}_{\rho}(y, u; \gamma) &= J(y, u) + \frac{\rho}{2} \{(K + \alpha M)y - \beta M_v \mathbf{u}\}^T \{(K + \alpha M)y - \beta M_v \mathbf{u}\} \\ &\quad + \gamma^T \{(K + \alpha M)y - \beta M_v \mathbf{u}\}, \end{aligned} \quad (2.3)$$

这里 $\rho > 0$ 为惩罚参数。因此, 求解问题(2.1)-(2.2)的交替乘子方向法算法表述如下:

算法1: 求解上述问题的交替乘子方向法算法(2.1)-(2.2).

步骤1: 设初值 $\{\mathbf{u}^0, \gamma^0\}$ 在 $U \times Y$ 中.

步骤2: 满足 $k \geq 0$, $\{\mathbf{u}^k, \gamma^k\} \rightarrow \mathbf{y}^{k+1} \rightarrow \mathbf{u}^{k+1} \rightarrow \gamma^{k+1}$ 通过求解与计算

$$\begin{cases} \mathbf{y}^{k+1} = \arg \min_{\mathbf{y} \in Y} \mathcal{L}_{\rho}(\mathbf{u}^k, \mathbf{y}; \gamma^k), & (2.4) \end{cases}$$

$$\begin{cases} \mathbf{u}^{k+1} = \arg \min_{\mathbf{u} \in U} \mathcal{L}_{\rho}(\mathbf{u}, \mathbf{y}^{k+1}; \gamma^k), & (2.5) \end{cases}$$

$$\begin{cases} \gamma^{k+1} = \gamma^k + \rho \{(K + \alpha M)\mathbf{y} - \beta M_v \mathbf{u}\}. & (2.6) \end{cases}$$

步骤3: 停止

3. 收敛性分析

3.1. 初步结果

本章将证明算法1的收敛性。为便于后续推导, 我们将等式 $(K + \alpha M)y = \beta M_v u$ 改写为另一种形式, 即 $y = Au$, 其中 $A := (K + \alpha M)^{-1} \beta M_v$ 。由此, 原问题可重新表述为:

$$\begin{cases} \min_{(\mathbf{u}, \mathbf{y}) \in U \times Y} [J(\mathbf{u}) + \delta_{[a, b]}(\mathbf{y})] \\ s.t. \quad \mathbf{y} = A\mathbf{u} \end{cases} \quad (3.1)$$

式子(2.3)可改写为

$$\mathcal{L}_\rho(\mathbf{y}, \mathbf{y}; \gamma) = J(u) + \delta_{[a,b]}(\mathbf{y}) + \frac{\rho}{2}(\mathbf{A}\mathbf{u} - \mathbf{y})^T(\mathbf{A}\mathbf{u} - \mathbf{y}) + \gamma^T(\mathbf{A}\mathbf{u} - \mathbf{y}). \quad (3.2)$$

利用(2.4)-(2.6)对应的凸性与次梯度条件, 我们可得如下方程组:

$$\begin{cases} \delta_{[a,b]}(\boldsymbol{\xi}) - \delta_{[a,b]}(\mathbf{y}^{k+1}) - \{\gamma^k + \rho(\mathbf{A}\mathbf{u}^k - \mathbf{y}^{k+1})\}^T(\boldsymbol{\xi} - \mathbf{y}^{k+1}) \geq 0, & \forall \boldsymbol{\xi} \in Y, & (3.3) \\ \nabla J(\mathbf{u}^{k+1}) + \rho A^T(\mathbf{A}\mathbf{u}^{k+1} - \mathbf{y}^{k+1} + \frac{\gamma^k}{\rho}) = 0, & & (3.4) \\ \gamma^{k+1} = \gamma^k + \rho(\mathbf{A}\mathbf{u}^{k+1} - \mathbf{y}^{k+1}). & & (3.5) \end{cases}$$

$$r_k := \mathbf{A}\mathbf{u}^k - \mathbf{y}^k. \quad (3.6)$$

将式(3.6)代入(3.3)-(3.5), 经变换后可得如下新的方程组:

$$\begin{cases} \delta_{[a,b]}(\boldsymbol{\xi}) - \delta_{[a,b]}(\mathbf{y}^{k+1}) - \{\gamma^{k+1} - \rho A(\mathbf{u}^{k+1} - \mathbf{u}^k)\}^T(\boldsymbol{\xi} - \mathbf{y}^{k+1}) \geq 0, \forall \boldsymbol{\xi} \in Y, & (3.7) \\ \nabla J(\mathbf{u}^{k+1}) + A^T \gamma^{k+1} = 0, & (3.8) \\ \rho r_{k+1} = \gamma^{k+1} - \gamma^k. & (3.9) \end{cases}$$

我们引入向量值函数 \mathbf{w} , 其定义为:

$$\mathbf{w} = \begin{pmatrix} \mathbf{u} \\ \mathbf{y} \end{pmatrix} \in \mathbb{R}^{2N}$$

我们定义 $\theta: \mathbb{R}^{2N} \rightarrow \mathbb{R} \cup +\infty$ 满足

$$\theta(\mathbf{w}) = J(\mathbf{u}) + \delta_{[a,b]}(\mathbf{y}), \quad \forall \mathbf{w} = (\mathbf{u}, \mathbf{y})^T \in \mathbb{R}^{2N}. \quad (3.10)$$

显然地, 函数 $J(u)$ 和 $\delta_{[a,b]}(y)$ 是凸的, θ 具有下半连续性. 于是有,

$$\partial\theta(\mathbf{w}) := \left\{ \mathbf{s} = \begin{pmatrix} \nabla J(\mathbf{u}) \\ \mathbf{q} \end{pmatrix} \mid \mathbf{q} \in \partial\delta_{[a,b]}(\mathbf{y}) \right\} \quad (3.11)$$

$$\mathbf{s}^T \hat{\mathbf{w}} = \{\nabla J(\mathbf{u})\}^T \hat{\mathbf{u}} + \mathbf{q}^T \hat{\mathbf{y}}, \quad \forall \mathbf{q} \in \partial\delta_{[a,b]}(\mathbf{y}), \hat{\mathbf{w}} \in \mathbb{R}^{2N}.$$

对任意方向 $\mathbf{q} \in \partial\delta_{[a,b]}(\mathbf{y})$, 在 \mathbf{w} 处关于次梯度的Bregman距离记为 $D_{\partial\theta(\mathbf{w})}\theta(\hat{\mathbf{w}}, \mathbf{w}; \mathbf{s})$, 其定义为:

$$D_{\partial\theta(\mathbf{w})}\theta(\hat{\mathbf{w}}, \mathbf{w}; \mathbf{s}) := \theta(\hat{\mathbf{w}}) - \theta(\mathbf{w}) - \mathbf{s}^T(\hat{\mathbf{w}} - \mathbf{w}), \quad \forall \hat{\mathbf{w}} \in \mathbb{R}^{2N} \quad (3.12)$$

因为 J 是严格凸的

$$(\nabla J(\mathbf{w}) - \nabla J(\mathbf{v}))^T(\mathbf{w} - \mathbf{v}) \geq \|\mathbf{w} - \mathbf{v}\|^2, \quad \forall \mathbf{v}, \mathbf{w} \in U. \quad (3.13)$$

此外, 由于 $\delta_{[a,b]}$ 也是凸函数, 因此Bregman距离满足如下不等式:

$$D_{\partial\theta(\mathbf{w})}\theta(\hat{\mathbf{w}}, \mathbf{w}; \mathbf{s}) \geq \frac{1}{2}\|\mathbf{u} - \hat{\mathbf{u}}\|^2, \quad \forall \mathbf{w}, \hat{\mathbf{w}} \in \mathbb{R}^{2N}. \quad (3.14)$$

根据式(3.8)、(3.7)和(3.11), \mathbf{s} 可重新表示为:

$$\mathbf{s}^k := \begin{pmatrix} -A^T \gamma^k \\ \gamma^k + \rho A(u^{k-1} - u^k) \end{pmatrix}. \quad (3.15)$$

3.2. 算法1全局收敛性

引理3.1. 设 $\{\mathbf{u}^k\}$ 和 $\{\mathbf{y}^k\}$ 是由算法1生成的序列,且 r_k 满足(3.6)的定义.对于任意的 $k \geq 0$ 满足下式

$$\begin{aligned} \rho A^T r_{k+1} &= \nabla J(\mathbf{u}^k) - \nabla J(\mathbf{u}^{k+1}), \\ \rho(r_{k+1})^T(\mathbf{y}^{k+1} - \mathbf{y}^k) &\geq \rho(A(\mathbf{u}^{k+1} - \mathbf{u}^k) - A(\mathbf{u}^k - \mathbf{u}^{k-1}))^T(\mathbf{y}^{k+1} - \mathbf{y}^k). \end{aligned}$$

引理3.2. 设 $E_k = \rho\|r_k\|^2 + \rho\|A(\mathbf{u}^k - \mathbf{u}^{k-1})\|^2$,则有

$$E_{k+1} \leq E_k - 2\|\mathbf{u}^{k+1} - \mathbf{u}^k\|^2. \quad (3.16)$$

由式(3.16)易知, E_k 是一个单调递减序列,且还满足如下性质:

$$\sum_{k=1}^{\infty} \|\mathbf{u}^{k+1} - \mathbf{u}^k\|^2 \leq \sum_{k=1}^{\infty} (E_k - E_{k+1}) = \frac{1}{2}(E_1 - E_{\infty})$$

$E_k > 0$ 并且始终单调递减,故 E_{∞} 是有界的,则

$$\sum_{k=1}^{\infty} \|\mathbf{u}^{k+1} - \mathbf{u}^k\|^2 < \infty \quad (3.17)$$

引理3.3. 设 $\{\mathbf{w}^k = \begin{pmatrix} \mathbf{u}^k \\ \mathbf{y}^k \end{pmatrix}\}$ 是算法1生成序列且 $\{D_{\partial\theta(\mathbf{w}^k)}\theta(\mathbf{w}^{k+1}, \mathbf{w}^k; \mathbf{s}^k)\}$ 满足(3.12)的定义.则

$$\sum_{k=1}^{\infty} (D_{\partial\theta(\mathbf{w}^k)}\theta(\mathbf{w}^{k+1}, \mathbf{w}^k; \mathbf{s}^k) + E_k) < \infty, \quad (3.18)$$

$$\sum_{k=1}^{\infty} \|A(\mathbf{u}^{k+1} - \mathbf{u}^k)\|^2 < \infty. \quad (3.19)$$

此外,序列 $\{\mathbf{u}^k\}$ 和 $\{\mathbf{y}^k\}$ 分别限制在 $L^2(\Omega)$ 和 $H^1(\Omega)$ 内.

由式3.19以及引理3.2中 E_k 的定义,可得:

$$\sum_{k=1}^{\infty} \|r_k\|^2 < \infty \quad \text{和} \quad r_k \rightarrow 0, \quad \text{则} \quad A\mathbf{u}^k - \mathbf{y}^k \rightarrow 0, \quad \text{随着} \quad k \rightarrow +\infty,$$

则有

$$\sum_{k=1}^{\infty} E_k < \infty. \quad (3.20)$$

引理3.4. 设 $\{\mathbf{w}^k = \begin{pmatrix} \mathbf{u}^k \\ \mathbf{y}^k \end{pmatrix}\}$ 由算法1生成且 $\{D_{\partial\theta(\mathbf{w}^k)}\theta(\hat{\mathbf{w}}, \mathbf{w}^k; \mathbf{s}^k)\}$ 在(3.12)中定义满足. 则当且仅当 $\hat{\mathbf{w}} = \begin{pmatrix} \hat{\mathbf{u}} \\ \hat{\mathbf{y}} \end{pmatrix}$ 是一个可行点, 有 $\{D_{\partial\theta(\mathbf{w}^k)}\theta(\hat{\mathbf{w}}, \mathbf{w}^k; \mathbf{s}^k)\}$ 收敛到 $k \rightarrow +\infty$.

引理3.5. 序列 $\{\mathbf{u}^n\}$ 和 $\{\mathbf{y}^n\}$ 分别是 $L^2(\Omega)$ 和 $H^1(\Omega)$ 的柯西序列.

根据引理3.5, $\mathbf{u}^n, \mathbf{y}^n$ 都是在完备空间 $L^2(\Omega)$ 和 $H^1(\Omega)$ 的Cauchy序列, 这意味着存在 $\bar{\mathbf{u}}, \bar{\mathbf{y}}$ 使得:

$$\mathbf{u}^n \rightarrow \bar{\mathbf{u}} \quad \text{and} \quad \mathbf{y}^n \rightarrow \bar{\mathbf{y}}, \quad \text{as } n \rightarrow +\infty \quad (3.21)$$

同时, 作为问题(3.1)的可行点, $\bar{\mathbf{w}} = \begin{pmatrix} \bar{\mathbf{u}} \\ \bar{\mathbf{y}} \end{pmatrix}$ 满足 $A\bar{\mathbf{u}} - \bar{\mathbf{y}} = 0$. 后文, [18]的证明技巧以及前面所给出的理论将被用以证明本文的收敛性.

定理3.6. 设 $\mathbf{w}^* = (\mathbf{u}, \mathbf{y})^T$ 是问题算法1的唯一解. 则随着 $n \rightarrow +\infty$, 有 $\mathbf{u}^n \rightarrow \mathbf{u}$ 和 $\mathbf{y}^n \rightarrow \mathbf{y}$. 同时随着 $n \rightarrow +\infty$, 有 $J(\mathbf{u}^n) + \delta(\mathbf{y}^n) \rightarrow J(\mathbf{u}) + \delta(\mathbf{y})$.

证明: 第一步, 我们将证明迭代序列收敛, 第二步再证明收敛点是最优点. 根据式(3.11)、式(3.8)、式(3.9)和式(3.6)处理下式:

$$\begin{aligned} & (\mathbf{s}^n - \mathbf{s}^m)^T (\mathbf{w}^n - \hat{\mathbf{w}}) \\ &= (\nabla J(\mathbf{u}^n) - \nabla J(\mathbf{u}^m))^T (\mathbf{u}^n - \hat{\mathbf{u}}) + (\mathbf{q}^n - \mathbf{q}^m)^T (\mathbf{y}^n - \hat{\mathbf{y}}) \\ &= \rho(A(\mathbf{u}^m - \mathbf{u}^{m-1}) - A(\mathbf{u}^n - \mathbf{u}^{n-1}), \mathbf{y}^n - \hat{\mathbf{y}}) - \rho \sum_{k=m}^{n-1} r_{k+1}^T r_k \end{aligned}$$

利用柯西-施瓦茨不等式以及 E_n 本身的单调递减性 ($k < n$), 可得如下估计:

$$\begin{aligned} & |(\mathbf{s}^n - \mathbf{s}^m)^T (\mathbf{w}^n - \hat{\mathbf{w}})| \\ & \leq \sum_{k=m}^{n-1} \frac{\rho}{2} (\|r_{k+1}\|^2 + \|r_n\|^2) + \rho |(A(\mathbf{u}^m - \mathbf{u}^{m-1}))^T (\mathbf{y}^n - \hat{\mathbf{y}})| + \rho |(A(\mathbf{u}^m - \mathbf{u}^{m-1}))^T (\mathbf{y}^n - \hat{\mathbf{y}})| \\ & \leq \frac{\rho}{2} \sum_{k=m+1}^n \|r_k\|^2 + \frac{n-m}{2} \rho \|r_n\|^2 + \rho \|\mathbf{y}^n - \hat{\mathbf{y}}\| (\|A(\mathbf{u}^m - \mathbf{u}^{m-1})\| + \|A(\mathbf{u}^m - \mathbf{u}^{m-1})\|) \\ & \leq \frac{1}{2} \sum_{k=m+1}^n E_k + \frac{n-m}{2} E_n + \rho \|\mathbf{y}^n - \hat{\mathbf{y}}\| (\|A(\mathbf{u}^m - \mathbf{u}^{m-1})\| + \|A(\mathbf{u}^m - \mathbf{u}^{m-1})\|) \\ & \leq \sum_{k=m+1}^n E_k + \rho \|\mathbf{y}^n - \hat{\mathbf{y}}\| (\|A(\mathbf{u}^m - \mathbf{u}^{m-1})\| + \|A(\mathbf{u}^m - \mathbf{u}^{m-1})\|) \end{aligned} \quad (3.22)$$

此外, 根据引理3.5和式(3.17)的结果, 将 $\hat{\mathbf{w}} = \bar{\mathbf{w}}$ 代入式(3.22), 可得如下收敛准则:

$$\rho\|\mathbf{y} - \bar{\mathbf{y}}\|(\|A(\mathbf{u}^m - \mathbf{u}^{m-1})\| + \|A(\mathbf{u}^n - \mathbf{u}^{n-1})\|) \rightarrow 0 \quad \text{as } n \rightarrow \infty$$

有以下估计式,

$$|(\mathbf{s}^n - \mathbf{s}^m)^T(\mathbf{w}^n - \bar{\mathbf{w}})| \leq \sum_{k=m+1}^{+\infty} E_k \quad \text{as } n \rightarrow \infty$$

三角不等式以及极限保不等式性,

$$\limsup_{n \rightarrow \infty} |(\mathbf{s}^n)^T(\mathbf{w}^n - \bar{\mathbf{w}})| \leq \sum_{k=m+1}^{+\infty} E_k \quad (3.23)$$

次梯度条件得到,

$$\theta(\mathbf{w}^n) \leq \theta(\bar{\mathbf{w}}) + (\mathbf{s}^n)^T(\mathbf{w}^n - \bar{\mathbf{w}}) \quad (3.24)$$

由于式(3.20)表明了 $\sum_{k=1}^n E_k$ 收敛即其余项满足当 $m \rightarrow +\infty$ 时 $\sum_{k=m+1}^n E_k \rightarrow 0$, 将该结论与式(3.23)和式(3.24)相结合, 可得:

$$\limsup_{n \rightarrow +\infty} \theta(\mathbf{w}^n) \leq \theta(\bar{\mathbf{w}}) \quad (3.25)$$

利用(3.21)和 θ 的下半连续性, 我们有

$$\theta(\bar{\mathbf{w}}) \leq \liminf_{n \rightarrow +\infty} \theta(\mathbf{w}^n) \quad (3.26)$$

结合(3.25)和(3.26)有

$$\theta(\bar{\mathbf{w}}) = \lim_{n \rightarrow +\infty} \theta(\mathbf{w}^n)$$

接下来我们将证明收敛点 $\bar{\mathbf{w}}$ 是最优点, 即对任意的 $\hat{\mathbf{w}}$, 都有 $\theta(\bar{\mathbf{w}}) \leq \theta(\hat{\mathbf{w}})$ 。基于前文分析、引理3.3、式(3.17)以及式(3.22)的结果, 可得如下收敛准则:

$$|(\mathbf{s}^n - \mathbf{s}^m)^T(\mathbf{w}^n - \hat{\mathbf{w}})| \rightarrow 0 \quad \text{as } n, m \rightarrow +\infty \quad (3.27)$$

极限的保号性, 对任意的 $\epsilon > 0$, 存在 $M > 0$, 使得任意的 $n > M$,

$$\rho|(\mathbf{u}^M - \mathbf{u}^{M-1})^T A^T(\mathbf{y}^n - \hat{\mathbf{y}})| < \epsilon, \quad (3.28)$$

$$|(\mathbf{s}^n - \mathbf{s}^M)^T(\mathbf{w}^n - \hat{\mathbf{w}})| < \epsilon, \quad (3.29)$$

次梯度条件和(3.15), (3.29), (3.28),

$$\begin{aligned} \theta(\mathbf{w}^n) &\leq \theta(\hat{\mathbf{w}}) + (\mathbf{s}^n)^T(\mathbf{w}^n - \hat{\mathbf{w}}) \\ &\leq \theta(\hat{\mathbf{w}}) + (\mathbf{s}^n - \mathbf{s}^M)^T(\mathbf{w}^n - \hat{\mathbf{w}}) + (\mathbf{s}^M)^T(\mathbf{w}^n - \hat{\mathbf{w}}) \\ &\leq \theta(\hat{\mathbf{w}}) + \epsilon + \mathbf{s}^T(\mathbf{w}^n - \hat{\mathbf{w}}) \end{aligned}$$

$$\begin{aligned}
&= \theta(\hat{\mathbf{w}}) + \epsilon - (A\gamma^M)^T(\mathbf{u}^n - \hat{\mathbf{u}}) + (\gamma^M - \rho A(\mathbf{u}^M - \mathbf{u}^{M-1}))^T(\mathbf{y}^n - \hat{\mathbf{y}}) \\
&= \theta(\hat{\mathbf{w}}) + \epsilon - (\gamma^M)^T(\mathbf{y}^n - \hat{\mathbf{y}} + r_n) + (\gamma^M - \rho A(\mathbf{u}^M - \mathbf{u}^{M-1}))^T(\mathbf{y}^n - \hat{\mathbf{y}}) \\
&\leq \theta(\hat{\mathbf{w}}) + 2\epsilon + \|\gamma^M\| \|r_n\|
\end{aligned}$$

此处，由于当 $\|r_n\| \rightarrow 0$ 时， $n \rightarrow +\infty$ ，由此可得：

$$\limsup_{n \rightarrow \infty} \theta(\mathbf{w}^n) \leq \theta(\hat{\mathbf{w}}) + 2\epsilon$$

与证明的前文部分类似，我们仍需利用下半连续性，已知 $\mathbf{w}^n \rightarrow \bar{\mathbf{w}}$ ，在此基础上建立 $\hat{\mathbf{w}}$ 与 $\bar{\mathbf{w}}$ 之间的关系，

$$\theta(\bar{\mathbf{w}}) \leq \liminf_{n \rightarrow +\infty} \theta(\mathbf{w}^n) \leq \limsup_{n \rightarrow +\infty} \theta(\mathbf{w}^n) \leq \theta(\hat{\mathbf{w}}) + 2\epsilon$$

于是有

$$\theta(\bar{\mathbf{w}}) \leq \theta(\hat{\mathbf{w}})$$

基于此，我们已证明可行点 $\bar{\mathbf{w}}$ 恰好是所求的最优点 \mathbf{w}^* ，即 $\bar{\mathbf{w}} = \mathbf{w}^*$ 且 $\theta(\bar{\mathbf{w}}) = \theta(\mathbf{w}^*)$ 。

证明完毕。 \square

3.3. 改进交替方向乘子法全局收敛性

交替方向乘子法的收敛速率对 ρ 的选取非常敏感。若 ρ 取值过大，会减慢子问题的更新速度；若 ρ 取值过小，则会影响算法对约束条件的满足度，两种情况都会导致收敛速度变慢。基于这一前提，本文采用自适应调整的 ρ 算子对算法1进行改进。此外，自适应能够显著提升算法的鲁棒性，降低算法对初始 ρ 还能显著提高算法的鲁棒性，降低算法对初始 ρ 取值的依赖。

这里我们给出算法2的全局收敛性，

算法2: 自适应交替乘子方向法求解(2.1)-(2.2)

步骤1: 设置初值 $j, \theta_1, \theta_2, \theta_3, \theta_4, \{\mathbf{u}^0, \gamma^0, \rho^0\}$ 满足 $U \times Y \times \mathbb{R}$ 。

步骤2: 对于 $k \geq 0, \{\mathbf{u}^k, \gamma^k\} \rightarrow \mathbf{y}^{k+1} \rightarrow \mathbf{u}^{k+1} \rightarrow \gamma^{k+1}$ 计算如下式子并求解

$$\left\{ \begin{array}{l} \mathbf{y}^{k+1} = \arg \min_{\mathbf{y} \in Y} \mathcal{L}_{\rho^k}(\mathbf{u}^k, \mathbf{y}; \gamma^k), \end{array} \right. \quad (3.30)$$

$$\left\{ \begin{array}{l} \mathbf{u}^{k+1} = \arg \min_{\mathbf{u} \in U} \mathcal{L}_{\rho^k}(\mathbf{u}, \mathbf{y}^{k+1}; \gamma^k), \end{array} \right. \quad (3.31)$$

$$\left\{ \begin{array}{l} \gamma^{k+1} = \gamma^k + \rho^k \{(K + \alpha M)\mathbf{y} - \beta M_v \mathbf{u}\}. \end{array} \right. \quad (3.32)$$

$$\left\{ \begin{array}{l} \rho^{k+1} = \begin{cases} j * \rho & \theta_1 \leq r_k, \theta_3 \leq e_k, \\ \frac{\rho}{j} & \theta_2 > r_k, \theta_4 > e_k, \\ \rho & \text{else.} \end{cases} \end{array} \right. \quad (3.33)$$

步骤3: 停止

定理3.7. 设 $\mathbf{w}^* = (\mathbf{u}, \mathbf{y})^T$ 是算法2的唯一解。则随着 $n \rightarrow +\infty$, 有 $\mathbf{u}^n \rightarrow \mathbf{u}$ 和 $\mathbf{y}^n \rightarrow \mathbf{y}$ 。同时随着 $n \rightarrow +\infty$, 有 $J(\mathbf{u}^n) + \delta(\mathbf{y}^n) \rightarrow J(\mathbf{u}) + \delta(\mathbf{y})$ 。

定理3.6的证明与定理3.7的证明完全类似, 只需要将定理3.6中的不变罚因子 ρ 改成自适应变化罚因子 ρ^k 即可, 此处不再赘述。接下来我们将给出算法2对应的两类收敛速率证明。

3.4. 收敛速率分析

本节将介绍自适应交替方向乘子法的遍历收敛速率与非遍历收敛速率, 以分析该算法在求解凸椭圆最优控制问题时的 $O(1/n)$ 最坏情况收敛速率。基于上述假设与理论结果, 推导得到该算法的遍历收敛速率。此外, 还给出了非遍历收敛速率的不同证明过程。

将式(3.5)代入式(3.4), 经变形后可得:

$$\nabla J(\mathbf{u}^{k+1}) + A^T \gamma^{k+1} - \rho A^T A(\mathbf{u}^{k+1} - \mathbf{u}^k) = 0$$

对偶可行性条件式(3.8),

$$\nabla J(\mathbf{u}^{k+1}) + A^T \gamma^{k+1} = 0$$

两式子对比, 我们得到对偶残差 e_k

$$e_k = -\rho A^T A(\mathbf{u}^{k+1} - \mathbf{u}^k) = -\rho A^T (\mathbf{y}^{k+1} - \mathbf{y}^k)$$

原始残差 r_k 和对偶残差 e_k 都要满足

$$r_k, e_k \rightarrow 0 \quad \text{as } k \rightarrow \infty$$

通过将每次迭代中原残差 r_k 与对偶残差 e_k 的大小作为 ρ 的自适应调整准则, 我们提出如下自适应交替方向乘子法算法。接下来, 我们将分别给出遍历收敛速率与非遍历收敛速率的定理证明。

定理3.8. 若 $\{\mathbf{w}^k = \begin{pmatrix} \mathbf{u}^k \\ \mathbf{y}^k \end{pmatrix}\}$ 是算法2生成的迭代序列且 r_k 由(3.6)定义得到。则对任意满足 $n \geq 1$, 我们定义

$$\tilde{\mathbf{w}}_n = \frac{1}{n} \sum_{k=1}^n \mathbf{w}^k, \quad \text{and} \quad \tilde{r}_n = \frac{1}{n} \sum_{k=1}^n r_k.$$

于是最优值点 $\{\mathbf{w}^* = \begin{pmatrix} \mathbf{u} \\ \mathbf{y} \end{pmatrix}\}$, 满足

$$\theta(\tilde{\mathbf{w}}_n) - \theta(\mathbf{w}^*) + \frac{\rho}{2} \|\tilde{r}_n\|^2 \leq \frac{1}{n} \left(\frac{\rho}{2} \|A(\mathbf{u}^0 - \mathbf{u})\|^2 + \frac{1}{2\rho} \|\gamma^0\|^2 \right). \quad (3.34)$$

证明: $\theta(\mathbf{w})$ 的凸性且 $\mathbf{u}^k \in \partial\theta(\mathbf{w}^k)$, 我们有

$$\theta(\mathbf{w}^k) - \theta(\mathbf{w}^*) \leq (\mathbf{s}^k)^T (\mathbf{w}^k - \mathbf{w}^*) \quad (3.35)$$

根据(3.33)的定义, 我们假设 ρ 发生了 p 次增大, q 减小, o 不变, 即

$$\rho^k = j^{p-q} \rho^0, \quad p + q + o = k, \quad p, q, o \in \mathbb{N}^+$$

根据(3.6)-(3.8), 得到

$$\begin{aligned} (\mathbf{s}^k)^T(\mathbf{w}^k - \mathbf{w}^*) &= -(\gamma^k)^T A(\mathbf{u}^k - \mathbf{u}) + \{\gamma^k + \rho A(\mathbf{u}^{k-1} - \mathbf{u}^k)\}^T(\mathbf{y}^k - \mathbf{y}) \\ &= -(\gamma^k)^T(\mathbf{y}^k - \mathbf{y} + r_k) + \{\gamma^k + \rho A(\mathbf{u}^{k-1} - \mathbf{u}^k)\}^T(\mathbf{y}^k - \mathbf{y}) \\ &= -\rho\{A(\mathbf{u}^k - \mathbf{u}^{k-1})\}^T(\mathbf{y}^k - \mathbf{y}) - (\gamma^k)^T r_k. \end{aligned} \quad (3.36)$$

结合(3.34)和(3.36),

$$\theta(\mathbf{w}^k) - \theta(\mathbf{w}^*) \leq -\rho\{A(\mathbf{u}^k - \mathbf{u}^{k-1})\}^T(\mathbf{y}^k - \mathbf{y}) - (\gamma^k)^T r_k. \quad (3.37)$$

给出如下等式

$$(a - b)^T b = \frac{1}{2}\{\|a\|^2 - \|a - b\|^2 - \|b\|^2\}, \quad (3.38)$$

由引理3.1和(3.38)

$$\begin{aligned} & -\rho\{A(\mathbf{u}^k - \mathbf{u}^{k-1})\}^T(\mathbf{y}^k - \mathbf{y}) \\ &= \rho\{A(\mathbf{u}^k - \mathbf{u}^{k-1})\}^T r_k - \rho\{A(\mathbf{u}^k - \mathbf{u}^{k-1})\}^T A(\mathbf{u}^k - \mathbf{u}) \\ &= \frac{\rho}{2}\{\|A(\mathbf{u}^{k-1} - \mathbf{u})\|^2 - \|A(\mathbf{u}^k - \mathbf{u})\|^2 - \|A(\mathbf{u}^k - \mathbf{u}^{k-1})\|^2\}, \end{aligned} \quad (3.39)$$

上式变化取 $a = A(\mathbf{u}^{k-1} - \mathbf{u})$ 和 $b = A(\mathbf{u}^k - \mathbf{u})$ 得到。

同理, 根据(3.9)和(3.38)有

$$\begin{aligned} -(\gamma^k)^T r_k &= \frac{1}{\rho}(\gamma^k, \gamma^{k-1} - \gamma^k) = \frac{1}{2\rho}(\|\gamma^{k-1}\|^2 - \|\gamma^{k-1} - \gamma^k\|^2 - \|\gamma^k\|^2) \\ &= \frac{1}{2\rho}(\|\gamma^{k-1}\|^2 - \|\gamma^k\|^2) - \frac{\rho}{2}\|r_k\|^2. \end{aligned} \quad (3.40)$$

根据(3.37), (3.39)和(3.40), 我们得到

$$\theta(\mathbf{w}^k) - \theta(\mathbf{w}^*) + \frac{\rho}{2} \leq \frac{\rho}{2}\{\|A(\mathbf{u}^{k-1} - \mathbf{u})\|^2 - \|A(\mathbf{u}^k - \mathbf{u})\|^2\} + \frac{1}{2\rho}(\|\gamma^{k-1}\|^2 + \|\gamma^k\|^2). \quad (3.41)$$

由 θ 是凸性, 下式子满足

$$\begin{aligned} \theta(\tilde{\mathbf{w}}_n) &\leq \frac{1}{n} \sum_{k=1}^n \theta(\mathbf{w}^k), \\ \|\tilde{r}_n\|^2 &\leq \frac{1}{n} \sum_{k=1}^n \|r_k\|^2. \end{aligned}$$

对 $k = 1$ 到 $k = n$ 求和, 得到

$$\theta(\tilde{\mathbf{w}}_n) - \theta(\mathbf{w}^*) + \frac{\beta}{2} \|\tilde{r}_n\|^2 \leq \frac{1}{n} \left(\frac{\beta}{2} \|S(u^0 - u)\|^2 + \frac{1}{2\beta} \|\lambda^0\|^2 \right),$$

证明完毕. □

$\{\mathbf{w}^k = \begin{pmatrix} \mathbf{u}^k \\ \mathbf{y}^k \end{pmatrix}\}$ 是收敛到最优值点 $\{\mathbf{w}^* = \begin{pmatrix} \mathbf{u} \\ \mathbf{y} \end{pmatrix}\}$ 的迭代序列, 引入以下向量:

$$\xi^k := \begin{pmatrix} \mathbf{w}^k \\ \gamma^k \end{pmatrix} = \begin{pmatrix} \mathbf{u}^k \\ \mathbf{y}^k \\ \gamma^k \end{pmatrix} \in \mathbb{R}^{3N} \quad \text{and} \quad \xi^* := \begin{pmatrix} \mathbf{u} \\ \mathbf{y} \\ \gamma \end{pmatrix},$$

由定理3.6, $\mathbf{w}^k \rightarrow \mathbf{w}^*$ 随着 $k \rightarrow +\infty$, 则同样有

$$\xi^k \rightarrow \xi^* \quad \text{as} \quad k \rightarrow \infty$$

再次由(3.33)的定义,

$$\rho^k = j^{p-q} \rho^0, \quad p + q + o = k, \quad p, q, o \in \mathbb{N}^+$$

于是, 引入以下向量 $\tilde{\xi}^k$:

$$\tilde{\xi}^k = \begin{pmatrix} \mathbf{u}^{k+1} \\ \mathbf{y}^{k+1} \\ \gamma^k + j^{p-q} \rho^0 (\mathbf{y}^{k+1} - \mathbf{A}\mathbf{u}^k) \end{pmatrix}.$$

定理3.9. 假设通过(2.4)-(2.6)计算得到序列 $\{\tilde{\xi}^k\}_{k=1}^{\infty}$ 的每一个分量. 那么

$$\lim_{k \rightarrow \infty} \left\{ k \cdot \min_{i=1,2,\dots,k} \{F(\tilde{\xi}^i)\} \right\} = 0. \quad (3.42)$$

证明: 基于该问题的一阶最优性条件, 定义函数 $F(\tilde{\xi}^k) : \mathbb{R}^{3N} \rightarrow [0, +\infty)$,

$$\begin{aligned} F(\tilde{\xi}^k) &= \|G - A[\gamma^k + j^{p-q} \rho^0 (\mathbf{y}^{k+1} - \mathbf{A}\mathbf{u}^k)]\|^2 + \|\mathbf{y}^{k+1} - \mathbf{A}\mathbf{u}^k\|^2 \\ &\quad + \text{dist}^2(0, \delta_{[a,b]}(\mathbf{y}^{k+1}) + [\gamma^k + j^{p-q} \rho^0 (\mathbf{y}^{k+1} - \mathbf{A}\mathbf{u}^k)]), \end{aligned} \quad (3.43)$$

其中 $G = \lambda_{\Omega} AM_v(\mathbf{A}\mathbf{u}^{k+1} - y_{\Omega}) + \lambda_{\Gamma} AM_g(\mathbf{A}\mathbf{u}^{k+1} - y_{\Gamma}) + \lambda M_v \mathbf{u}^{k+1}$.

$$0 \in \text{dist}^2(0, \delta_{[a,b]}(\mathbf{y}^{k+1}) + [\gamma^k + j^{p-q} \rho^0 (\mathbf{y}^{k+1} - \mathbf{A}\mathbf{u}^k)]),$$

$$G - A[\gamma^k + j^{p-q} \rho^0 (\mathbf{y}^{k+1} - \mathbf{A}\mathbf{u}^{k+1})] = 0,$$

随后将上述两个等式代入 $F(\tilde{\xi}^k)$ 结合式(3.6)与引理3.2中 E_k 的定义, 令 $\rho = \rho^0$, 则存在 $C > 0$, 使得

$$\begin{aligned} F(\tilde{\xi}^k) &= \|j^{p-q}\rho^0 A^2(\mathbf{u}^{k+1} - \mathbf{u}^k)\|^2 + \|r_{k+1}\|^2 \\ &= j^{2(p-q)}\rho^0 \|A\|^2 \rho^0 \|A(\mathbf{u}^{k+1} - \mathbf{u}^k)\|^2 + \|r_{k+1}\|^2 \\ &\leq C j^{2(p-q)} \left\{ \rho^0 \|A(\mathbf{u}^{k+1} - \mathbf{u}^k)\|^2 + \rho^0 \|r_{k+1}\|^2 \right\} \\ &\leq C \frac{E_k j^{2K}}{j^{2q}} \leq C \frac{E_k}{j^{2q}}. \end{aligned}$$

根据式(3.20), 级数 $\sum_{k=0}^{\infty} E_k$ 是收敛且分母 j^{2q} 随着 k 增大至少以线性速率增长, 因此 $\frac{E_k}{j^{2q}}$ 至少以指数衰减速率递减。由此可知正项级数 $\sum_{k=0}^{\infty} \frac{E_k}{j^{2q}}$ 收敛, 进而可推出正项级数 $\sum_{k=0}^{\infty} F(\tilde{\xi}^k)$ 是收敛的。再根据文献[19]中的引理6.1, 式(3.42)成立。

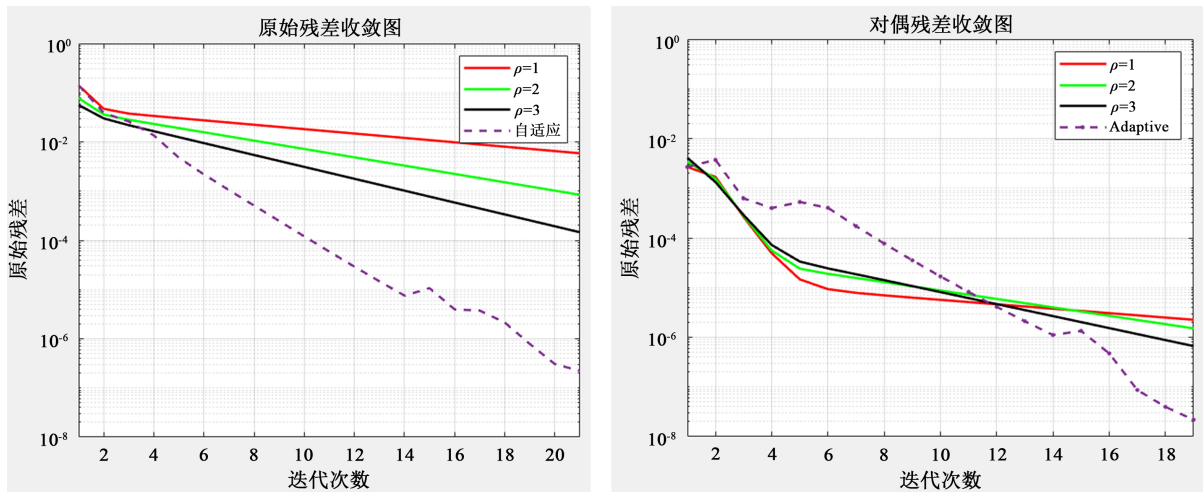


Figure 1. Convergence curve

图 1. 收敛分析图

4. 数值实验

本节我们将通过数值算例验证所提出的改进交替方向乘子法的有效性。采用有限元法对问题进行离散化, 单元数设置为64, 所有算例均通过MATLAB软件求解。结果主要通过图表形式呈现: 其中图(图 1)用于展示不同算法的收敛速率对比, 表则分别列出迭代次数iter, 控制变量 u 的 L^2 误差, 状态变量 y 与目标状态 y_{Ω} (在区域 Ω 内)和 y_{Γ} (在边界 Γ 内)的误差, 以及损失泛函 $J(y, u)$ 的取值。问题的停止容差设置为 10^{-6} , 以下为问题的具体设置,

$$\begin{aligned} \lambda &= 0.2, \quad \lambda_{\Omega} = \lambda_{\Gamma} = \beta = \alpha = 1, \\ y_{\Omega} &= \sin(\pi x_2) \sin(\pi x_1), \quad y_{\Gamma} = y_{\Omega}|_{\Gamma}. \end{aligned}$$

从原残差与对偶残差的收敛曲线可以看出, 在其余参数保持不变的情况下, 自适应交替方向乘

Table 1. ADMM solution results for different ρ **表 1.** 不同下 ρ 的交替乘子方向法求解结果

ρ	<i>iter</i>	$\ y - y_\Omega\ _{L^2(\Omega)}$	$\ y - y_\Gamma\ _{L^2(\Omega)}$	$J(y, u)$	$\ u\ _{L^2(\Omega)}$
0.5	-	-	-	-	-
1	43	5.63376e-01	2.10195e-01	1.92040e-01	3.35454e-01
2	32	5.66067e-01	2.14567e-01	1.94598e-01	3.37089e-01
3	23	5.66075e-01	2.14551e-01	1.94602e-01	3.37136e-01
adaptive	18	5.66136e-01	2.14428e-01	1.94654e-01	3.37782e-01

子法所需迭代次数最少，即可达到收敛容差要求。表 1 中的数据也印证了这一点：对于 $\rho = 0.5$ 时，在我们的 10^{-6} 标准下无法达到收敛，自适应调整的 ρ 仅需 18 次迭代即可收敛，远少于固定 ρ 所需的迭代次数，充分体现了其通过动态调整参数平衡原可行性与对偶可行性的优势。

5. 总结

本文针对带罗宾边界条件的椭圆型最优控制问题，提出一种带自适应罚参数 ρ 的交替方向乘子法。理论层面，无需依赖拉格朗日乘子的存在性，严格证明了算法的全局收敛性；同时建立了遍历与非遍历两类收敛速率，证明该算法在凸问题框架下最坏收敛复杂度为 $O(1/n)$ 。数值层面，通过有限元离散将问题转化为有限维优化模型，引入自适应 ρ 更新策略，可根据原残差与对偶残差的动态变化自动调整罚参数。实验结果表明，自适应策略在保证求解精度的同时显著减少了迭代次数，且状态拟合误差与目标函数值和固定参数方法的结果高度一致，验证了所提算法的鲁棒性与高效性。

参考文献

- [1] Mehdaoui, M., Lacitignola, D. and Tilioua, M. (2025) State-Constrained Optimal Control of a Coupled Quasilinear Parabolic System Modeling Economic Growth in the Presence of Technological Progress. *Applied Mathematics & Optimization*, **91**, Article No. 14. <https://doi.org/10.1007/s00245-024-10214-6>
- [2] Enaet Hossain, K., Liang, D. and Zhu, H. (2024) Developing PDE-Constrained Optimal Control of Multicomponent Contamination Flows in Porous Media. *Computers & Mathematics with Applications*, **176**, 224-243. <https://doi.org/10.1016/j.camwa.2024.10.033>
- [3] Li, S., Yang, L. and Gao, Z. (2020) Distributed Optimal Control for Multiple High-Speed Train Movement: An Alternating Direction Method of Multipliers. *Automatica*, **112**, Article 108646. <https://doi.org/10.1016/j.automatica.2019.108646>
- [4] Pierer von Esch, M., Landgraf, D., Steffel, M., Völz, A. and Graichen, K. (2024) Distributed

- Stochastic Optimal Control of Nonlinear Systems Based on ADMM. *IEEE Control Systems Letters*, **8**, 424-429. <https://doi.org/10.1109/lcsys.2024.3393411>
- [5] Chen, X., Song, X., Chen, Z. and Yu, B. (2020) A Multi-Level ADMM Algorithm for Elliptic PDE-Constrained Optimization Problems. *Computational and Applied Mathematics*, **39**, Article No. 331. <https://doi.org/10.1007/s40314-020-01379-1>
- [6] Li, X., Sun, D. and Toh, K. (2016) A Schur Complement Based Semi-Proximal ADMM for Convex Quadratic Conic Programming and Extensions. *Mathematical Programming*, **155**, 333-373. <https://doi.org/10.1007/s10107-014-0850-5>
- [7] Chen, C., He, B., Ye, Y. and Yuan, X. (2016) The Direct Extension of ADMM for Multi-Block Convex Minimization Problems Is Not Necessarily Convergent. *Mathematical Programming*, **155**, 57-79. <https://doi.org/10.1007/s10107-014-0826-5>
- [8] Song, Y., Yuan, X. and Yue, H. (2024) The ADMM-PINNs Algorithmic Framework for Non-smooth PDE-Constrained Optimization: A Deep Learning Approach. *SIAM Journal on Scientific Computing*, **46**, C659-C687. <https://doi.org/10.1137/23m1566935>
- [9] Song, Y., Wang, Z. and Zuazua, E. (2025) FedADMM-InSa: An Inexact and Self-Adaptive ADMM for Federated Learning. *Neural Networks*, **181**, Article 106772. <https://doi.org/10.1016/j.neunet.2024.106772>
- [10] Chen, L., He, S. and Fan, X. (2025) Cooperative Optimization of Shared Energy Storage in Integrated Energy Systems Using Adaptive ADMM and Nash Bargaining. *Journal of Energy Storage*, **134**, Article 118148. <https://doi.org/10.1016/j.est.2025.118148>
- [11] Chen, P., Ye, Y., Wang, H., Bu, S., Tang, Y. and Strbac, G. (2024) Holistic Coordination of Transactive Energy and Carbon Emission Right Trading for Heterogenous Networked Multi-Energy Microgrids: A Fully Distributed Adaptive Consensus ADMM Approach. *Sustainable Energy Technologies and Assessments*, **64**, Article 103729. <https://doi.org/10.1016/j.seta.2024.103729>
- [12] Langer, U., Löscher, R., Steinbach, O. and Yang, H. (2024) An Adaptive Finite Element Method for Distributed Elliptic Optimal Control Problems with Variable Energy Regularization. *Computers & Mathematics with Applications*, **160**, 1-14. <https://doi.org/10.1016/j.camwa.2024.02.006>
- [13] Banz, L., Hintermüller, M. and Schröder, A. (2025) *hp*-Finite Elements for Elliptic Optimal Control Problems with Control Constraints. *Computers & Mathematics with Applications*, **196**, 288-311. <https://doi.org/10.1016/j.camwa.2025.07.030>
- [14] Brenner, S.C., Sung, L. and Zhang, Y. (2019) C^0 Interior Penalty Methods for an Elliptic State-Constrained Optimal Control Problem with Neumann Boundary Condition. *Journal of*

-
- Computational and Applied Mathematics*, **350**, 212-232.
<https://doi.org/10.1016/j.cam.2018.10.015>
- [15] Yang, C. and Sun, T. (2023) Second-Order Time Discretization for Reaction Coefficient Estimation of Bilinear Parabolic Optimization Problem with Neumann Boundary Conditions. *Computers & Mathematics with Applications*, **140**, 211-224.
<https://doi.org/10.1016/j.camwa.2023.04.016>
- [16] Wang, T. and Zhou, Z. (2025) Multilevel Correction Finite Element Method for Semilinear Elliptic Optimal Control Problems with Bilinear Robin Boundary Conditions. *Communications in Nonlinear Science and Numerical Simulation*, **151**, Article 109052.
<https://doi.org/10.1016/j.cnsns.2025.109052>
- [17] Chrysafinos, K., Gunzburger, M.D. and Hou, L.S. (2006) Semidiscrete Approximations of Optimal Robin Boundary Control Problems Constrained by Semilinear Parabolic PDE. *Journal of Mathematical Analysis and Applications*, **323**, 891-912.
<https://doi.org/10.1016/j.jmaa.2005.10.053>
- [18] Glowinski, R., Song, Y. and Yuan, X. (2020) An ADMM Numerical Approach to Linear Parabolic State Constrained Optimal Control Problems. *Numerische Mathematik*, **144**, 931-966. <https://doi.org/10.1007/s00211-020-01104-4>
- [19] Chen, L., Sun, D. and Toh, K. (2017) An Efficient Inexact Symmetric Gauss-Seidel Based Majorized ADMM for High-Dimensional Convex Composite Conic Programming. *Mathematical Programming*, **161**, 237-270. <https://doi.org/10.1007/s10107-016-1007-5>