

控制与决策

Control and Decision

基于奇异摄动的复杂工业过程强化学习运行优化控制

周林娜, 蔡天赐, 赵建国, 熊梦辉, 杨春雨

引用本文:

周林娜, 蔡天赐, 赵建国, 等. 基于奇异摄动的复杂工业过程强化学习运行优化控制[J]. *控制与决策*, 2025, 40(5): 1581-1589.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2024.0700>

您可能感兴趣的其他文章

Articles you may be interested in

混合柯西变异和均匀分布的蝗虫优化算法

Hybrid Cauchy mutation and uniform distribution of grasshopper optimization algorithm

控制与决策. 2021, 36(7): 1558-1568 <https://doi.org/10.13195/j.kzyjc.2019.1609>

输入约束不确定系统的点对点迭代学习控制与优化

Point-to-point iterative learning control and optimization for uncertain systems with constrained input

控制与决策. 2021, 36(6): 1435-1441 <https://doi.org/10.13195/j.kzyjc.2019.0908>

基于复杂昂贵仿真的体系效能多目标优化

Complex and expensive simulation based multi-objective optimization to system-of-system effectiveness

控制与决策. 2021, 36(3): 589-598 <https://doi.org/10.13195/j.kzyjc.2019.0844>

基于深度强化学习与迭代贪婪的流水车间调度优化

Scheduling optimization for flow-shop based on deep reinforcement learning and iterative greedy method

控制与决策. 2021, 36(11): 2609-2617 <https://doi.org/10.13195/j.kzyjc.2020.0608>

基于强化学习的多目标车辆跟随决策算法

Multi-objective vehicle following decision algorithm based on reinforcement learning

控制与决策. 2021, 36(10): 2497-2503 <https://doi.org/10.13195/j.kzyjc.2020.0426>

基于奇异摄动的复杂工业过程强化学习运行优化控制

周林娜, 蔡天赐, 赵建国[†], 熊梦辉, 杨春雨

(中国矿业大学 信息与控制工程学院, 江苏 徐州 221116)

摘要: 复杂工业过程往往存在快、慢耦合动态特性, 应用常规的串级设计思路难以保障整体运行优化性能, 而现有整体设计运行优化控制的方法容易发生“维数灾”和“病态数值”问题. 鉴于此, 聚焦一类设备单元和运行过程模型参数皆未知的复杂工业过程, 提出一种基于奇异摄动的快慢并行强化学习组合非串级运行优化控制方法. 首先, 引入收敛因子将复杂工业过程运行优化控制问题建模为双时间尺度系统的非串级最优控制问题; 其次, 使用奇异摄动理论将原始最优问题分解为降阶快子系统最优调节和慢子系统最优设定值跟踪问题; 然后, 在强化学习的框架下, 设计数据驱动的迭代算法学习快、慢子系统最优控制器, 进而构造不依赖于系统模型的组合最优控制策略. 与现有方法相比, 所提出运行优化算法不仅能解决设备单元模型参数未知问题, 还能实现对期望运行指标的零误差渐近跟踪. 最后, 通过赤铁矿混合选别浓密过程对比实验验证所提出方法的有效性和先进性.

关键词: 复杂工业过程; 奇异摄动理论; 运行优化控制; 强化学习

中图分类号: TP29

文献标志码: A

DOI: 10.13195/j.kzyjc.2024.0700

引用格式: 周林娜, 蔡天赐, 赵建国, 等. 基于奇异摄动的复杂工业过程强化学习运行优化控制 [J]. 控制与决策, 2025, 40(5): 1581-1589.

Reinforcement learning-based optimal operational control of complex industrial processes using singular perturbations

ZHOU Lin-na, CAI Tian-ci, ZHAO Jian-guo[†], XIONG Meng-hui, YANG Chun-yu

(School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221116, China)

Abstract: Complex industrial processes are often subject to fast and slow coupling dynamic characteristics. Using the traditional cascade design approach cannot ensure the optimal operation performance of the whole system, while the existing integrated design methods tend to result in high dimensionality and ill-conditioned numerics. For a class of complex industrial processes where both unit device and operational process are unknown, we propose a parallel fast and slow reinforcement learning-based composite non-cascade control approach using singular perturbations. Firstly, the optimal operational control problem of complex industrial processes is modeled as a non-cascade optimal control problem of a two-time-scale system introducing a convergence factor. Secondly, the original optimal control problem is decomposed into an optimal regulation problem of a fast subsystem and an optimal set-point tracking problem of a slow subsystem via singular perturbation theory. Thirdly, we design a data-driven iterative algorithm to learn the optimal controllers of both fast and slow subsystems in the framework of reinforcement learning, and then we construct a composite optimal control strategy that is independent of model parameters. Compared to the existing methods, the proposed approach does not require knowledge of model parameters of device layer and achieves zero-error asymptotic tracking of desired operational index. Finally, we validate the effectiveness and superiority of the proposed method through experimental results conducted on mixed separation thickening process.

Keywords: complex industrial processes; singular perturbation theory; optimal operational control; reinforcement learning

收稿日期: 2024-06-12; 录用日期: 2024-10-26.

基金项目: 国家自然科学基金项目 (62073327, 62273350, 62403467, 62403466); 江苏省基础研究计划自然科学基金项目 (BK20241635); 国家资助博士后研究人员计划项目 (GZB20240827); 江苏省卓越博士后计划项目 (2024ZB604, 2024ZB835); 江苏省研究生科研与实践创新计划项目 (KYCX24_2774); 中国矿业大学研究生创新计划项目 (2024WLJCRCZL139).

责任编辑: 刘向杰.

[†]通信作者. E-mail: jianguo Zhao@cumt.edu.cn.

0 引言

复杂工业过程具有快慢动态耦合、模型参数不确定等特性,其优化与自动控制研究逐渐成为控制科学与工程领域的难点和热点^[1-2].工业过程的自动化不仅要确保过程被控变量准确跟踪设定值,更重要的是要在整个生产过程中控制与产品质量、效率和资源消耗相关的指标,确保这些运行指标始终保持在目标范围内,从而实现过程运行的优化控制.工业过程的运行优化控制旨在全面考虑生产线的运行状况,在确保设备单元安全运行的前提下控制运行过程,提高产品加工的质量和效率,同时尽量减少资源消耗.因此,工业过程运行优化控制具有理论和实际应用价值,也引起了学术界和工业界的广泛关注^[3-5].

工业过程运行优化控制问题的传统串级控制框架是一个双环结构.外环利用设定值优化器,根据期望的运行指标 r^* 确定设定值 $w(t)$.内环利用设定值跟踪控制器,确保运行过程的运行指标 $r(t)$ 能够跟踪期望指标或保持在目标范围内^[4-6].这种串级控制结构的优势在于能够分别设计内环跟踪器和外环优化器^[7-8].另一方面,考虑到工业生产过程变量复杂、规模大、机理不确定,难以获取系统的准确模型参数^[9],一些学者尝试将自适应控制和工业过程相结合.文献^[10]面向动态未知的工业过程提出了一种基于神经网络补偿和多模型切换的自适应控制方法,通过整合内环控制系统和外环优化模型,有效地改善了控制算法的瞬态性能.文献^[11]提出了一种基于补偿信号的铁矿石混合选别浓密过程双速率自适应控制方法,使磨矿过程中的各项指标在参数未知变化的情况下保持在目标值.然而,这些工业过程自适应控制方法缺乏优化能力.

强化学习(RL)作为机器学习的重要分支,已经用于求解动态未知系统的最优控制问题^[12-14],也在工业过程运行优化控制应用中取得显著成效.同时,考虑到工业过程往往表现出多时间尺度特性^[4],许多学者已经在工业过程的运行优化问题中引入了提升技术.文献^[15]提出了一种双速率数据驱动算法,该算法基于提升技术和RL,旨在使运行指标保持在目标范围内.文献^[16]针对密介质分离煤炭清洗过程,提出了一种基于RL的非线性PI控制器,并结合提升技术,用于工业运行优化控制.文献^[17]将基于提升技术的零和博弈理论应用到未知的工业过程,使用离线RL算法来解决运行优化问题.文献^[18]使用提升技术将设备单元基于模型的模型预测控制与运行层基于数据的RL相结合来解决大规模工业过程

运行优化问题.但是,由于现代工业过程的复杂化与大型化,使用提升技术的工业过程运行优化方法往往会导致其控制系统阶数过高,容易在控制设计时发生“维数灾”和“病态数值”问题^[4].

为了避免这些问题,学者们开始研究工业过程的非串级控制方案,用于实现设备单元与运行过程的同步调节^[3-6].非串级控制的特点是建模工业过程为双时间尺度系统,其中设备层动态相比运行层动态快得多^[5-6,15],通过设计适当的控制输入 $u(t)$ 使运行指标 $r(t)$ 保持在其期望的目标值 r^* .奇异摄动理论是双时间尺度系统建模的有力工具,在现代工程领域得到了广泛的应用^[19].这类系统中存在一个小参数,用来表示快、慢状态变量之间的时间尺度差异^[20].根据奇异摄动理论,双时间尺度系统可以解耦为独立的快、慢子系统,然后通过降阶的快、慢子系统设计原系统的组合控制器^[21].文献^[3]利用奇异摄动理论分别设计快、慢子系统的最优控制问题,将原始最优问题分解为慢子问题和快子问题,并结合Q-learning算法,成功实现了运行指标的跟踪并有效降低了控制器复杂度.文献^[22]针对双时间尺度耦合的工业过程,提出了一种离线RL算法,将原系统跟踪问题分解为慢子系统的跟踪问题和快子系统的镇定问题,实现了理想的运行指标跟踪.文献^[5]使用奇异摄动理论设计了双时间尺度大型工业过程的分布式控制算法,实现了设定值近似最优跟踪控制.在文献^[4]中,针对包含多个设备单元的工业过程,提出了一种由数据和模型共同驱动的非串级组合分散控制策略.这一方法有效地减少了控制器的通讯量和计算负荷,并有效抑制了外部干扰.然而,工业过程的非串级运行优化控制仍面临许多亟待解决的科学难题.首先,现有工业过程运行优化控制成果大多假设设备单元动态精确已知或可辨识,这在大规模复杂工业过程中难以实现;其次,使用折扣代价函数的工业过程运行优化算法存在跟踪误差,不能实现零误差跟踪.

本文将综合考虑复杂工业过程的特点,结合RL技术和奇异摄动理论,针对一类快、慢耦合且模型未知的工业过程,提出一种完全数据驱动的非串级组合运行优化控制方法.主要工作概括如下:

1) 区别于已有的工业过程运行优化非串级控制方法^[3-4,6,18,22],本文基于带有收敛因子的代价函数,设计完全无模型的工业过程RL非串级运行优化控制器,解决了现存方法要求设备单元模型参数已知的问题.

2) 不同于现存基于折扣代价函数的最优跟踪控

制方法^[3-5, 15-17, 22], 本文通过修订代价函数优化控制输入微分的加权项, 移除代价函数中的折扣因子, 实现了期望运行指标的零误差渐近跟踪。

1 工业过程非串级运行优化控制问题描述

本节首先将工业过程建模成一类快、慢动态未知的双时间尺度系统; 其次用基于带有收敛因子的代价函数表述工业过程非串级运行优化控制问题; 最后根据最优控制理论, 给出运行优化控制问题的全阶最优解。

1.1 工业过程双时间尺度建模

考虑文献 [4] 中的复杂工业过程, 设备层模型动态为

$$\begin{cases} \dot{x}_1(t) = \varepsilon^{-1} A_{11} x_1(t) + \varepsilon^{-1} B_1 u(t), \\ y(t) = C_1 x_1(t). \end{cases} \quad (1)$$

其中: $\varepsilon > 0$ 是时间尺度参数; $x_1 \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, $y \in \mathbb{R}^q$ 分别表示系统状态、控制输入和被控输出; $A_{11} \in \mathbb{R}^{n \times n}$, $B_1 \in \mathbb{R}^{n \times m}$, $C_1 \in \mathbb{R}^{q \times n}$ 是适当维数的常数矩阵。

运行层模型动态为

$$\begin{cases} \dot{x}_0(t) = A_{00} x_0(t) + A_{01} y(t), \\ r(t) = C_0 x_0(t). \end{cases} \quad (2)$$

其中: $x_0 \in \mathbb{R}^{n_0}$ 是运行过程的系统状态; $r(t) \in \mathbb{R}^{q_0}$ 为运行指标; $A_{00} \in \mathbb{R}^{n_0 \times n_0}$, $A_{01} \in \mathbb{R}^{n_0 \times q}$, $C_0 \in \mathbb{R}^{q_0 \times n_0}$ 是适当维数的矩阵。为了后续控制设计, 作如下假设。

假设 1 矩阵 A_{11} 是非奇异的。

假设 2 时间尺度参数 ε 是已知的。

假设 3 矩阵 A_{11} , B_1 , C_1 , A_{00} , A_{01} 和 C_0 是未知的。

注 1 时间尺度参数 ε 描述工业过程的设备层与运行层状态变量之间的快慢分离程度。实际上, ε 往往为设备单元中的时间常数、质量或电容^[21]。对于许多实际系统而言, 它是已知的或名义值是已知的^[20, 23]。因此, 假设 2 对于很多工业过程来说是合理的^[4]。文献 [3-4, 6, 18, 22] 的工作假设了设备单元模型参数 A_{11} 、 B_1 和 C_1 是精确已知的, 无法适用于设备层动态不确定的复杂工业过程。本文在这些基础上进一步考虑设备单元动态矩阵也是未知的, 进而设计不依赖于设备层模型信息的运行优化控制器。

1.2 非串级运行优化控制问题

工业过程的运行优化控制需要优化代价函数并实现运行指标设定值的跟踪控制, 这里期望运行指标为常值。运行指标 $r(t)$ 反映了操作过程中的执行情况, 它应该跟踪目标值 r^* , 通常跟踪误差越小, 运行性能越好。

设计运行优化控制器, 即设备单元输入 u 最小化如下代价函数:

$$J = \int_0^\infty e^{2\alpha\tau} [\delta^T Q \delta + v^T R v] d\tau. \quad (3)$$

其中: $\delta = r - r^*$, $\dot{u} = v$, $Q > 0$ 和 $R > 0$ 是对称正定矩阵, $\alpha > 0$ 是收敛因子。

注 2 当使用标准最优控制问题定义运行优化控制问题时, 有界控制输入 u 将导致代价函数 J 无界。一些研究者尝试使用折扣因子 γ 来保证 J 的有界性, 但这种方法会出现跟踪误差^[24]。本文在代价函数中选择控制微分代替控制本身, 进而在不使用折扣因子 γ 的情况下保证 J 的有界性同时实现零误差渐近跟踪。此外, 本文引入收敛因子 α , 有助于克服组合最优控制策略设计时依赖设备层模型参数的问题, 且收敛因子 α 的引入并不影响代价函数 J 的有界性。

1.3 运行优化问题的全阶最优解

为了求解工业过程运行优化控制的最优反馈控制器, 将运行优化控制问题 (3) 表述为增广系统的最优调节器, 然后给出基于全阶模型的最优解。

令 $\zeta = [\dot{x}_0^T \quad \delta^T \quad \dot{x}_1^T]^T$, 可将式 (1) 和 (2) 写成如下增广系统:

$$\begin{aligned} \dot{\zeta} &= \begin{bmatrix} A_{00} & 0 & A_{01} C_1 \\ C_0 & 0 & 0 \\ 0 & 0 & \varepsilon^{-1} A_{11} \end{bmatrix} \zeta + \begin{bmatrix} 0 \\ 0 \\ \varepsilon^{-1} B_1 \end{bmatrix} v \equiv \\ & A\zeta + Bv. \end{aligned} \quad (4)$$

根据式 (4), 将代价函数 (3) 改写为

$$J = \int_0^\infty e^{2\alpha\tau} [\zeta^T Q_c \zeta + v^T R v] d\tau, \quad (5)$$

其中 $Q_c = \text{diag}\{0, Q, 0\}$ 。

根据最优控制理论^[25-26], 由式 (4) 和 (5) 构成的标准线性二次型调节问题的最优控制输入为

$$v_{\text{opt}} = K\zeta \equiv K_1 \dot{x}_0 + K_2 \delta + K_3 \dot{x}_1. \quad (6)$$

通过变换可得运行优化控制问题的全阶最优解为

$$u_{\text{opt}} = K_1 x_0 + K_2 z + K_3 x_1, \quad (7)$$

其中: $z = \int_0^t \delta d\tau$; $K = -R^{-1} B^T P$, P 满足如下代数 Riccati 方程:

$$A^T P + PA + Q_c - PBR^{-1} B^T P + 2\alpha P = 0. \quad (8)$$

2 基于奇异摄动的运行优化控制问题求解

令 $z_1 = [\dot{x}_0^T \quad \delta^T]^T$, $z_2 = \dot{x}_1$, 可将增广系统 (4) 改写成如下形式:

$$\begin{aligned} \dot{z}_1 &= \begin{bmatrix} A_{00} & 0 \\ C_0 & 0 \end{bmatrix} z_1 + \begin{bmatrix} A_{01} C_1 \\ 0 \end{bmatrix} z_2, \\ \varepsilon \dot{z}_2 &= A_{11} z_2 + B_1 v. \end{aligned} \quad (9)$$

式 (9) 是一个标准的双时间尺度系统, 其中 z_1 为慢状

态变量, z_2 为快状态变量. 因此, 直接设计全阶最优解 (7) 可能会面临“维数灾”和“病态数值”问题^[21]. 为了克服这些数值计算上的困难, 本节将基于奇异摄动理论对双时间尺度系统 (9) 进行快、慢子系统分解, 分别设计降阶子系统控制器进而给出运行优化控制问题的组合求解方法.

2.1 慢子系统最优控制

根据奇异摄动理论可知, 双时间尺度系统 (9) 中快状态 z_2 达到了准稳态的情况下, 可以忽略快变量的影响. 令式 (9) 中的 $\varepsilon = 0$, 忽略系统中的快尺度动态^[21], 则快状态 z_2 在慢尺度下可表示为

$$0 = A_{11}z_2^s(t) + B_1v^s(t). \quad (10)$$

在假设 1 下, 可将式 (10) 写成如下形式:

$$z_2^s(t) = -A_{11}^{-1}B_1v^s(t). \quad (11)$$

将式 (11) 代入 (9), 可得慢子系统模型

$$\dot{z}_1^s = \begin{bmatrix} A_{00} & 0 \\ C_0 & 0 \end{bmatrix} z_1^s + \begin{bmatrix} H_s \\ 0 \end{bmatrix} \equiv \bar{A}z_1^s + \bar{B}v^s. \quad (12)$$

其中: $z_1^s = [(\dot{x}_0^s)^T (\delta^s)^T]^T$, $H_s = -A_{01}C_1A_{11}^{-1}B_1$.

为了保证慢子系统最优问题解的存在性, 作如下常规假设.

假设 4 (A_{00}, H_s) 是可控的, (A_{00}, C_0) 是可观的.

依据经典奇异摄动理论, 代价函数 (5) 下慢子系统 (12) 对应的代价函数为

$$J_s = \int_0^\infty e^{2\alpha\tau} ((z_1^s)^T Q_s z_1^s + (v^s)^T R v^s) d\tau, \quad (13)$$

其中 $Q_s = [0 \ 1]^T Q [0 \ 1]$.

下面给出最小化代价函数 (13) 的可稳解存在且唯一的条件.

引理 1^[6] 在假设 4 下, 如果 A_{00} , H_s 和 C_0 满足如下条件:

$$\text{rank} \begin{bmatrix} A_{00} & H_s \\ C_0 & 0 \end{bmatrix} = n_0 + q_0, \quad (14)$$

则矩阵 (\bar{A}, \bar{B}) 是可控的, 矩阵 $(\bar{A}, \sqrt{Q_s})$ 是可观的.

在引理 1 下, 根据最优控制理论可得慢子系统 (12) 最小化代价函数 (13) 的控制输入为

$$v_{\text{opt}}^s = K_s z_1^s \equiv K_1^s \dot{x}_0^s + K_2^s \delta_s, \quad (15)$$

其中 $K_s = -R^{-1}\bar{B}^T P_s$, P_s 可由如下代数 Riccati 方程得到:

$$\bar{A}^T P_s + P_s \bar{A} + Q_s - P_s \bar{B} R^{-1} \bar{B}^T P_s + 2\alpha P_s = 0. \quad (16)$$

2.2 快子系统最优控制

定义快变量和系统输入存在如下等式:

$$\begin{aligned} z_2^f(t) &= z_2(t) - z_2^s(t), \\ v^f(t) &= v(t) - v^s(t). \end{aligned} \quad (17)$$

根据奇异摄动理论^[25], 在快时间尺度下, 可认为系统 (9) 的慢动态满足 $\varepsilon \dot{z}_2^s(t) \approx 0$. 通过式 (17) 和 (11), 有

$$\begin{aligned} \varepsilon \dot{z}_2(t) &= \varepsilon (\dot{z}_2^s(t) + \dot{z}_2^f(t)) = \\ &A_{11}(z_2^s(t) + z_2^f(t)) + B_1(v^s(t) + v^f(t)) = \\ &A_{11}z_2^f(t) + B_1v^f(t), \end{aligned}$$

可得到快子系统模型

$$\varepsilon \dot{z}_2^f(t) = A_{11}z_2^f(t) + B_1v^f(t). \quad (18)$$

相似于慢子系统, 代价函数 (5) 下快子系统 (18) 对应的代价函数为

$$J_f = \int_0^\infty e^{2\alpha\tau} (v^f)^T R v^f d\tau. \quad (19)$$

为了保证快子系统最优问题解的存在性, 作如下常规假设.

假设 5 (A_{11}, B_1) 是可控的.

假设 6 $\alpha > \max\{0, \max\{\text{Re}(-\lambda(A_{11}))\}\}$.

在假设 5 和假设 6 下, 根据最优控制理论^[25] 可得快子系统 (18) 最小化代价函数 (19) 的控制输入为

$$v_{\text{opt}}^f = K_f z_2^f, \quad (20)$$

其中 $K_f = -R^{-1}B_1^T P_f$, P_f 满足代数 Riccati 方程

$$P_f A_{11} + A_{11} P_f - P_f B_1 R^{-1} B_1^T P_f + 2\alpha P_f = 0. \quad (21)$$

2.3 组合最优控制

根据双时间尺度系统的最优控制理论, 可构造如下组合最优控制策略:

$$v_{\text{opt}}^c = v_{\text{opt}}^s + v_{\text{opt}}^f, \quad (22)$$

进而求解运行优化控制问题.

引理 2^[21] 如果闭环子系统 (12) 和 (18) 是渐近稳定的, 则 z_1 是系统 (9) 的慢状态, 存在 $\varepsilon^* > 0$, 对所有的 $\varepsilon \in (0, \varepsilon^*]$, 存在

$$z_1(t) = z_1^s(t) + o(\varepsilon).$$

由式 (11) 和 (17), 可得

$$z_2^f = z_2 + A_{11}^{-1}B_1v^s. \quad (23)$$

通过引理 2, 将组合控制器 (22) 改写为

$$\begin{aligned} v_{\text{opt}}^c &= K_s z_1^s + K_f z_2^f = \\ &(I + K_f A_{11}^{-1} B_1) K_s z_1 + K_f z_2. \end{aligned} \quad (24)$$

因为 $K_s = [K_1^s, K_2^s]$ 和 $v = \dot{u}$, 代入式 (24) 可得

$$\begin{aligned} u_{\text{opt}}^c &= K_f x_1 + (K_f A_{11}^{-1} B_1 + I) K_1^s x_0 + \\ &(K_f A_{11}^{-1} B_1 + I) K_2^s z \equiv K_c \eta. \end{aligned} \quad (25)$$

其中: $\eta = [x_1^T \ x_0^T \ z^T]^T$, u_{opt}^c 是所求组合运行优化控制策略.

定理 1 复杂工业过程 (1) 和 (2) 在控制输入 (25) 作用下, 期望运行指标的跟踪误差 $\delta(t) = r(t) - r^*$

渐近收敛到零.

证明 将式 (15) 代入 (12), 慢子系统可重写为

$$\dot{z}_1^s = A_{sc} [(x_0^s)^T \quad \delta^T]^T,$$

其中 $A_{sc} = \begin{bmatrix} A_{00} + H_s K_1^s & H_s K_2^s \\ C_0 & 0 \end{bmatrix}$. 因式 (15) 是式 (12) 的一个最优控制输入, 可知闭环矩阵 $A_{sc} = \bar{A} + \bar{B}K_s$ 的所有特征值在左半平面^[26]. 因此闭环慢子系统 (12) 渐近稳定.

将式 (21) 左侧同时乘 P_f^{-1} , 重写为如以下形式:

$$A_{11} - B_1 R^{-1} B_1^T P_f = P_f^{-1} (-A_{11}^T - 2\alpha I) P_f. \quad (26)$$

由式 (26) 可知, $A_{11} - B_1 R^{-1} B_1^T P_f$ 和 $-A_{11}^T - 2\alpha I$ 是相似的. 故可得

$$\text{Re}(\lambda(A_{11} - B_1 R^{-1} B_1^T P_f)) = \text{Re}(-\lambda(A_{11})) - 2\alpha.$$

当满足假设 6 时, $A_{11} + B_1 K^f$ 是赫尔维茨矩阵, 可知快子系统 (18) 渐近稳定.

根据奇异摄动理论, 双时间尺度系统 (9) 渐近稳定, 即在控制输入 (25) 作用下, 其状态变量即跟踪误差 $\delta(t) = r(t) - r^*$ 渐近收敛到零. \square

3 基于强化学习的组合最优控制

由式 (25) 可知, 实现工业过程组合最优控制设计需要已知 K_1^s 、 K_2^s 、 K_f 和 $K_f A_{11}^{-1} B_1$ 的信息. 传统获得 K_1^s 、 K_2^s 、 K_f 的方法需要基于模型求解非线性代数 Riccati 方程 (16)、(21). 但是, 由于工业过程的复杂性导致其机理不明确, 传统基于模型的方法很难在实际中应用. 本文将在 RL 的框架下, 提出一种不依赖模型参数 A_{11} 、 B_1 、 C_1 、 A_{00} 、 B_0 、 C_0 的控制方法, 基于工业过程数据的迭代算法来学习组合最优控制策略 (25).

3.1 强化学习算法符号定义

对于矩阵 $B = [b_1, \dots, b_m] \in \mathbb{R}^{n \times m}$, 矩阵 $H = [h_{ij}] \in \mathbb{R}^{n \times n}$ 和向量 $w = [w_i] \in \mathbb{R}^n$, 定义 $\text{vec}(B) = [b_1^T, \dots, b_m^T]^T \in \mathbb{R}^{nm}$, $\text{ves}(H) = [h_{11}, 2h_{12}, \dots, 2h_{1n}, 2h_{22}, \dots, 2h_{2n}, \dots, h_{nn}]^T \in \mathbb{R}^{0.5n(n+1)}$, $\text{vev}(w) = [w_1^2, w_1 w_2, \dots, w_1 w_n, w_2^2, \dots, w_2 w_n, \dots, w_n^2]^T \in \mathbb{R}^{0.5n(n+1)}$.

对于任意向量 β , ϑ 和一个充分大的数 $h > 0$, 进一步定义

$$\Gamma_\beta = [\text{vev}(\bar{\beta}(t_1)) - \text{vev}(\bar{\beta}(t_0)), \text{vev}(\bar{\beta}(t_2)) - \text{vev}(\bar{\beta}(t_1)), \dots, \text{vev}(\bar{\beta}(t_h)) - \text{vev}(\bar{\beta}(t_{h-1}))]^T,$$

$$\Pi_{\beta\vartheta} = \left[\int_{t_0}^{t_1} \bar{\beta} \otimes \bar{\vartheta} d\tau, \int_{t_1}^{t_2} \bar{\beta} \otimes \bar{\vartheta} d\tau, \dots, \int_{t_{h-1}}^{t_h} \bar{\beta} \otimes \bar{\vartheta} d\tau \right]^T.$$

其中: $\bar{\beta} = e^{\alpha t} \beta$, $\bar{\vartheta} = e^{\alpha t} \vartheta$, \otimes 代表克罗内克积, $0 \leq$

$t_0 < t_1 < \dots < t_{h-1} < t_h$ 是采样时刻.

3.2 慢子系统强化学习最优控制设计

本节将利用 RL 技术, 设计数据驱动的慢子系统最优控制增益 $K_s = [K_1^s, K_2^s]$ 的求解方法.

受基于模型的 Kleinman 算法启发^[27], 代数 Riccati 方程 (16) 可通过如下迭代算法求解:

$$(\bar{A} + \alpha I + \bar{B}K_s^k)^T P_s^k + P_s^k (\bar{A} + \alpha I + \bar{B}K_s^k) - Q_s - (K_s^k)^T R K_s^k. \quad (27)$$

其中: $K_s^{k+1} = -R^{-1} \bar{B}^T P_s^k$, $\bar{A} + \alpha I + \bar{B}K_s^0$ 是赫尔维茨的.

慢子系统 (12) 的运行数据含有 x_0^s 的微分数据, 在学习过程中不易于收集. 为此, 定义新的状态变量 $\varsigma = [(x_0^s)^T, z^T]^T$. 由式 (12) 可得慢辅助系统

$$\dot{\varsigma} = \begin{bmatrix} A_{00} & 0 \\ C_0 & 0 \end{bmatrix} \varsigma + \begin{bmatrix} -A_{01} C_1 A_{11}^{-1} B_1 \\ 0 \end{bmatrix} u^s + \begin{bmatrix} 0 \\ -r^* \end{bmatrix} \equiv \bar{A} \varsigma + \bar{B} u^s + \bar{D}, \quad (28)$$

其中 $\dot{u}^s = v^s$.

将 $u^s = K_s^k \varsigma + u_{se}$ 代入式 (28) 中, 其中 u_{se} 是探索噪声, 可得闭环慢辅助系统

$$\dot{\varsigma} = (\bar{A} + \bar{B}K_s^k) \varsigma + \bar{B}(u^s - K_s^k \varsigma) + \bar{D}. \quad (29)$$

带有收敛因子的代数 Riccati 方程 (16) 可由如下无模型 Bellman 方程迭代求解:

$$\begin{aligned} & e^{2\alpha(t+\delta t)} \varsigma^T(t+\delta t) P_s^k \varsigma(t+\delta t) - \\ & e^{2\alpha(t+\delta t)} \varsigma^T(t) P_s^k \varsigma(t) = \\ & - \int_t^{t+\delta t} e^{2\alpha\tau} (\varsigma^T Q_s \varsigma + (K_s^k \varsigma)^T R K_s^k \varsigma) d\tau - \\ & 2 \int_t^{t+\delta t} e^{2\alpha\tau} (u^s - K_s^k \varsigma)^T R K_s^{k+1} \varsigma d\tau + \\ & 2 \int_t^{t+\delta t} e^{2\alpha\tau} \bar{D}^T P_s^k \varsigma d\tau, \end{aligned} \quad (30)$$

其中 δt 是采样间隔. 通过克罗内克积运算, 重写积分方程 (30) 为如下矩阵方程:

$$\mathcal{X}_s^k \begin{bmatrix} \text{ves}(P_s^k) \\ \text{vec}(K_s^{k+1}) \\ \text{vec}(\bar{D}^T P_s^k) \end{bmatrix} = \mathcal{Y}_s^k. \quad (31)$$

其中: $\mathcal{X}_s^k = [\Gamma_\varsigma, -2(\Pi_{\varsigma\varsigma}(I \otimes (K_s^{k+1})^T R) - \Pi_{\varsigma u^s}(I \otimes R)), -2\Pi_{\varsigma I}]$, $\mathcal{Y}_s^k = -\Pi_{\varsigma\varsigma} \text{vec}(Q_s + (K_s^k)^T R K_s^k)$.

使用矩阵方程 (31) 可计算 P_s^k 、 K_s^{k+1} , 且不需要任何系统的知识. 但是, ς 中的 x_0^s 不能直接在原系统中收集. 类比引理 2, 可在学习过程中使用 $\varsigma_c = [(x_0)^T, z^T]^T$ 代替 ς . 因此, 进一步将式 (31) 改写为

$$\mathcal{X}^k \begin{bmatrix} \text{ves}(P_s^k) \\ \text{vec}(K_s^{k+1}) \\ \text{vec}(\bar{D}^T P_s^k) \end{bmatrix} = \mathcal{Y}^k. \quad (32)$$

其中

$$\begin{aligned}\mathcal{X}^k &= [\Gamma_{sc}, -2(\Pi_{scsc}(I \otimes (K_s^{k+1})^T R_s) - \\ &\quad \Pi_{scus}(I \otimes R_s)), -2\Pi_{scI}], \\ \mathcal{Y}^k &= -\Pi_{scsc} \text{vev}(Q_s + (K_s^k)^T R_s K_s^k).\end{aligned}$$

3.3 快子系统强化学习最优控制设计

利用 RL 技术, 设计数据驱动的快子系统最优控制增益求解方法, 同时学习 K_f 和 $K_f A_{11}^{-1} B_1$. 代数 Riccati 方程 (21) 可以用如下迭代算法求解:

$$\begin{aligned}(A_{11} + \alpha I + B_1 K_f^j)^T P_f^j + \\ P_f^j (A_{11} + \alpha I + B_1 K_f^j) = \\ - (K_f^j)^T R K_f^j.\end{aligned}\quad (33)$$

其中: $K_f^{j+1} = -R_f B_1^T P_f^j$, $A_{11} + \alpha I + B_1 K_f^j$ 是赫尔维茨的.

考虑行为策略 $u^f = K_f^j x_1^f + u_{fe}$. 其中: $\dot{u}^f = v^f$, u_{fe} 是探索噪声. 基于此构造快辅助系统

$$\varepsilon \dot{x}_1^f = (A_{11} + B_1 K_f^j) x_1^f + B_1 (u^f - K_f^j x_1^f). \quad (34)$$

由式 (33) 和 (34) 可得, 带有收敛因子的代数 Riccati 方程 (21) 可由如下无模型 Bellman 方程迭代求解

$$\begin{aligned}\varepsilon e^{2\alpha(t+\delta t)} (x_1^f)^T (t + \delta t) P_f^j x_1^f (t + \delta t) - \\ \varepsilon e^{2\alpha(t+\delta t)} (x_1^f)^T (t) P_f^j x_1^f (t) = \\ - \int_t^{t+\delta t} e^{2a\tau} (x_1^f)^T (K_f^j)^T R K_f^j x_1^f d\tau - \\ 2 \int_t^{t+\delta t} e^{2a\tau} (u^f - K_f^j x_1^f)^T R K_f^{j+1} x_1^f d\tau.\end{aligned}\quad (35)$$

参照式 (11), 令

$$x_1^f = x_1 + A_{11}^{-1} B_1 u^s \triangleq \Lambda \phi.$$

其中: $\Lambda = [I, A_{11}^{-1} B_1]$, $\phi = [x_1^T, (u^s)^T]^T$. 通过上述公式, 将式 (35) 重写为

$$\begin{aligned}\varepsilon e^{2\alpha(t+\delta t)} \phi^T (t + \delta t) \bar{P}_f^j \phi (t + \delta t) - \\ \varepsilon e^{2\alpha(t+\delta t)} \phi^T (t) \bar{P}_f^j \phi (t) = \\ - \int_t^{t+\delta t} e^{2a\tau} \phi^T (\bar{K}_f^j)^T R \bar{K}_f^j \phi d\tau - \\ 2 \int_t^{t+\delta t} e^{2a\tau} (u^f - \bar{K}_f^j \phi)^T R \bar{K}_f^{j+1} \phi d\tau.\end{aligned}\quad (36)$$

其中: $\bar{P}_f^j = \Lambda^T P_f^j \Lambda$, $\bar{K}_f^j = K_f^j \Lambda$.

通过克罗内克积运算和积分方程 (36), 得到如下矩阵方程

$$\mathcal{Z}^j \begin{bmatrix} \text{vec}(\bar{P}_f^j) \\ \text{vec}(\bar{K}_f^{j+1}) \end{bmatrix} = \mathcal{W}^j. \quad (37)$$

其中

$$\begin{aligned}\mathcal{Z}^j &= [\varepsilon \Gamma_\phi, -2\Pi_{\phi\phi}(I \otimes (\bar{K}_f^j)^T R_f) + \\ &\quad 2\Pi_{\phi u^f}(I \otimes R)], \\ \mathcal{W}^j &= -\Pi_{\phi\phi} \text{vev}((\bar{K}_f^j)^T R \bar{K}_f^j).\end{aligned}$$

下面引理给出了矩阵方程 (32) 和 (37) 的解唯一存在的条件.

引理 3^[25] 对于任意 $k, j \in \mathbb{N}$, 存在 $h^* \in \mathbb{N}$, 对于所有的 $h > h^*$, 如果满足

$$\begin{aligned}\text{rank}([\Pi_{\phi\phi}, \Pi_{\phi u^f}]) = \\ \frac{1}{2}(n_0 + q_0 + n)(n_0 + n_0 + 3n + 1),\end{aligned}\quad (38)$$

则矩阵 \mathcal{X}^k , \mathcal{Z}^j 对于任意 $k, j \in \mathbb{N}$ 都是列满秩的.

注 3 本文所设计的组合最优控制器 (25) 含有积分项 z , 使控制器可抑制常值或分段常值的干扰, 保持运行指标的零稳态误差跟踪. 在实践中, 这些干扰可能源于模型线性化误差、环境的变动、原料波动, 或是生产设备的老化等因素^[6]. 因此, 本文提出的方法对于复杂工业过程具有实际的工程应用价值.

4 仿真实验

以赤铁矿混合选别浓密过程 (MSTP)^[22] 为例, 验证所提出基于奇异摄动的 RL 非串级运行优化控制方法的有效性, 考虑复杂工况及扰动并通过与文献 [4,22] 进行对比, 验证本文所提出算法的先进性.

MSTP 是一种典型的双时间尺度工业系统, 其中作为设备单元浆体泵是快速过程, 而作为运行过程的底流浓度是慢速变化的^[3,4,22]. 本文将考虑以下由文献 [22] 提供的 MSTP 线性化动态模型

$$\begin{aligned}\dot{r}(t) &= -0.057r(t) + 0.055y(t), \\ \varepsilon \dot{y}(t) &= -0.068y(t) + 0.26u(t),\end{aligned}\quad (39)$$

时间尺度参数 $\varepsilon = 0.1$.

为了获得理想的浮选精矿品位和尾矿品位, 应将底流浓度 $r(t)$ 控制在目标范围内. 因此, 将期望的底流浓度设定为 $r^* = 33$, 选取代价函数 (3) 中的参数 $Q = 0.1I$, $R = 0.1I$, $\alpha = 0.07$.

通过选取初始增益矩阵 $K_s^0 = [3 \ 3]$, $\bar{K}_f^0 = [1 \ 1]$, 探测噪声 u^e 为不同频率随机正弦信号的和, 阈值选取 $\sigma_s = \sigma_f = 1 \times 10^{-6}$. 下面给出由本文提出的无模型算法得到的仿真结果. 通过迭代求解可得 $K_c = [0.0173 \ 3.2941 \ 1.2403]$, 仿真结果如图 1 ~ 图 4 所示.

由图 1 可看出快、慢子系统控制策略 K_s 、 \bar{K}_f 的收敛性; 由图 2 可知, 通过本文快慢并行组合强化学习算法, 产品指标 r 可以零误差地跟踪参考轨迹 r^* ; 由图 3 可知, 料浆流速 y 会收敛至其稳态值; 由图 4 可知, 底流浆体泵频率 u 会收敛至其稳态值. 仿真结果表明, 本文所提方法能够在模型参数未知下实现对具有双时间尺度特性的工业过程运行优化控制问题的无误差跟踪.

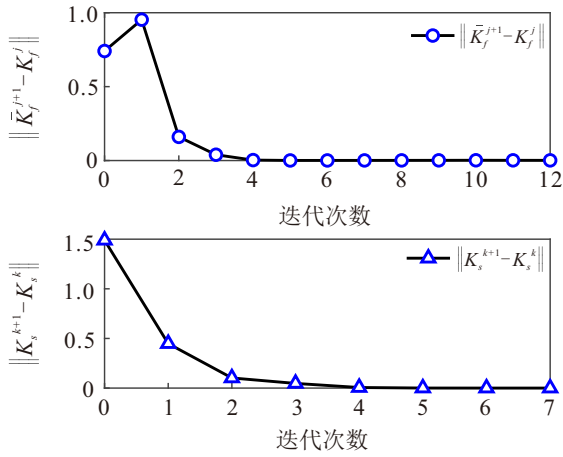


图1 学习过程中 K_s^k \bar{K}_f^j 和收敛曲线

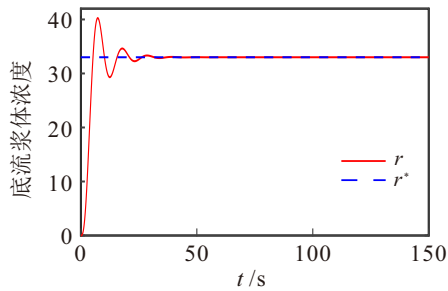


图2 底流浆体浓度轨迹

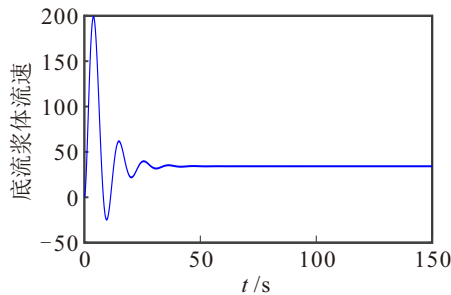


图3 底流浆体流速轨迹

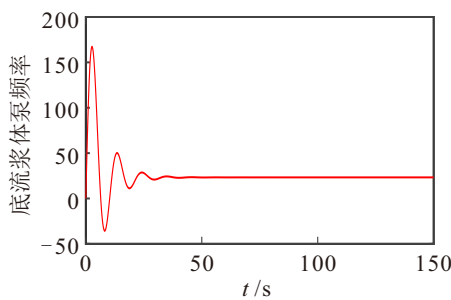


图4 底流浆体泵频率轨迹

在相同条件下, 通过求解代数 Riccati 方程, 可得快、慢子系统的最佳反馈增益矩阵分别为 $K_c = [0.0154 \ 3.2914 \ 1.1675]$. 仿真结果如图 5 所示, 显示了类似于图 2 的性能, 验证了本文所提出的无模型的性能与基于模型求解组合控制策略 (25) 的性能一致.

此外, 为了验证所提出工业过程运行优化控制方法在复杂工况下的运行情况, 在设备层加入如图 6

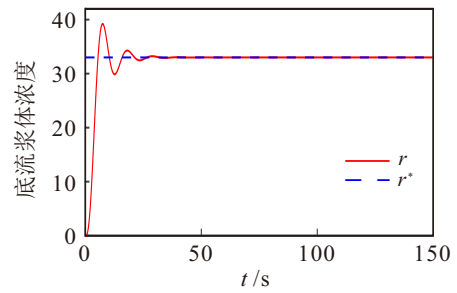


图5 通过求解代数 Riccati 方程得到的底流浓度的跟踪轨迹

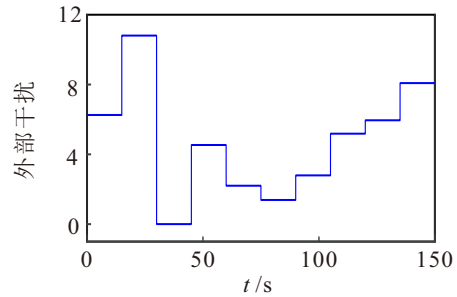


图6 外部扰动曲线

所示的干扰曲线, 用来模拟设备老化、环境变化等实际工况变化. 由图 7 可以看出, 尽管选别浓密过程受到干扰的影响, 但是在本文提出的控制器下, 底流浆体浓度依然稳定在其设定值. 该仿真结果表明, 本文提出的非串级优化控制器在扰动下具有良好的鲁棒性和抗干扰能力.

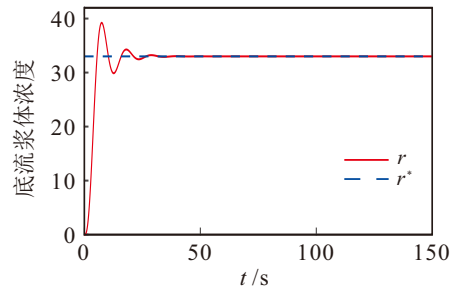


图7 扰动情况下底流浆体轨迹跟踪

为了体现本文快、慢并行组合 RL 算法的先进性, 将考虑与奇异摄动快、慢分解并假设设备单元系统已知的文献 [4] 和快子系统使用系统辨识方法的文献 [22] 进行对比, 仿真结果如图 8 所示. 其中: r 是

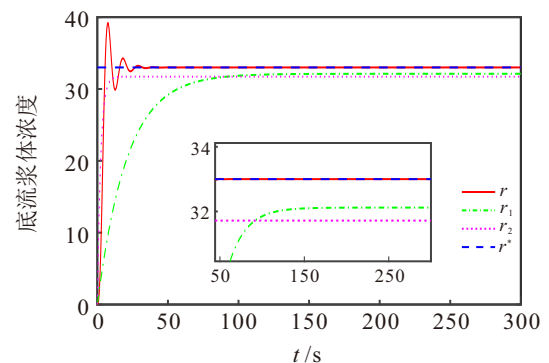


图8 底流浆体轨迹跟踪方法对比

本文所提出方法得到的底流浓度跟踪曲线, r_1 是文献 [22] 中方法得到的底流浓度跟踪曲线, r_2 是文献 [4] 中方法得到的底流浓度跟踪曲线。

为了具体评估控制性能, 下面使用绝对误差积分 (IAE) 和误差均方差 (MSE)^[5] 来评价控制效果:

$$\text{IAE} = \sum_{i=k^*}^{k^*+n} |r(i) - r^*(i)|,$$

$$\text{MSE} = \sqrt{\frac{1}{n} \sum_{i=k^*}^{k^*+n} |r(i) - r^*(i)|^2}.$$

其中: 选取 $k^* = 1500$, $n = 1500$, 即使用 150 s ~ 300 s 的数据进行对比, 对比结果如表 1 所示。

表1 评价指标对比

	IAE	MSE
本文算法	0	0
文献[4]算法	1925.4	1.28
文献[22]算法	1326.5	0.88

由图 8 和表 1 可以看出, 本文算法组合最优控制器能实现更理想的工业过程运行指标无误差渐近跟踪, 并且使用的系统知识更少, 无需设备层模型的具体参数. 主要原因是本文使用控制输入的微分来最小化代价函数, 避免了文献 [4] 和文献 [22] 使用折扣因子代价函数方法存在的跟踪误差问题; 其次, 本文使用基于收敛因子的 RL 算法也实现了完全无模型工业过程组合最优控制的设计问题。

5 结语

本文针对一类快慢耦合且模型参数完全未知的复杂工业过程, 提出了一种基于奇异摄动的快慢并行 RL 组合非串级运行优化控制方法. 具体优化控制方法包括非串级工业过程双时间尺度建模, 基于数据驱动的快、慢子系统控制器设计, 最终实现工业过程的整体优化和期望运行指标的渐近跟踪. 仿真结果表明, 所提出方法具有良好的自学习能力, 在不需要设备层和运行层模型参数的情况下, 能够使运行指标零误差跟踪目标值。

参考文献 (References)

- [1] 柴天佑, 程思宇, 李平, 等. 端边云协同的复杂工业过程运行控制智能系统[J]. 控制与决策, 2023, 38(8): 2051-2062.
(Chai T Y, Cheng S Y, Li P, et al. Intelligent system for operational control of complex industrial process based on end-edge-cloud collaboration[J]. Control and Decision, 2023, 38(8): 2051-2062.)
- [2] 柴天佑. 工业人工智能与工业互联网协同实现生产过程智能化及其未来展望[J]. 控制工程, 2023, 30(8):

1378-1388.

(Chai T Y. Industrial AI and industrial Internet collaboratively achieving production process intelligence and its future perspectives[J]. Control Engineering of China, 2023, 30(8): 1378-1388.)

- [3] Xue W Q, Fan J L, Lopez V G, et al. New methods for optimal operational control of industrial processes using reinforcement learning on two time scales[J]. IEEE Transactions on Industrial Informatics, 2020, 16(5): 3085-3099.
- [4] Zhao J G, Yang C Y, Dai W, et al. Reinforcement learning-based composite optimal operational control of industrial systems with multiple unit devices[J]. IEEE Transactions on Industrial Informatics, 2022, 18(2): 1091-1101.
- [5] Li J N, Nie H, Chai T Y, et al. Reinforcement learning for optimal tracking of large-scale systems with multitime scales[J]. Science China Information Sciences, 2023, 66(7): 170201.
- [6] 赵建国, 杨春雨. 复杂工业过程非串级双速率组合分散运行优化控制[J]. 自动化学报, 2023, 49(1): 172-184.
(Zhao J G, Yang C Y. Non-cascade dual-rate composite decentralized operational optimal control for complex industrial processes[J]. Acta Automatica Sinica, 2023, 49(1): 172-184.)
- [7] Meng S M, Huang W J, Yin X C, et al. Security-aware dynamic scheduling for real-time optimization in cloud-based industrial applications[J]. IEEE Transactions on Industrial Informatics, 2021, 17(6): 4219-4228.
- [8] Fan J L, Jiang Y, Chai T Y. MPC-based setpoint compensation with unreliable wireless communications and constrained operational conditions[J]. Neurocomputing, 2017, 270: 110-121.
- [9] 乔俊飞, 黄卫民, 丁海旭, 等. 复杂工业过程特征建模方法及应用研究[J]. 控制与决策, 2023, 38(8): 2063-2078.
(Qiao J F, Huang W M, Ding H X, et al. Research on feature modeling method for complex industrial process and its application[J]. Control and Decision, 2023, 38(8): 2063-2078.)
- [10] 富月, 杜琼. 一类工业运行过程多模型自适应控制方法[J]. 自动化学报, 2018, 44(7): 1250-1259.
(Fu Y, Du Q. Multi-model adaptive control method for a class of industrial operational processes[J]. Acta Automatica Sinica, 2018, 44(7): 1250-1259.)
- [11] Wang L Y, Jia Y, Chai T Y, et al. Dual-rate adaptive control for mixed separation thickening process using compensation signal based approach[J]. IEEE Transactions on Industrial Electronics, 2018, 65(4): 3621-3632.
- [12] 高伟男, 杨涛, 柴天佑. 基于自适应动态规划和梯度下降法的自适应最优输出调节[J]. 控制与决策, 2023, 38(8): 2425-2432.
(Gao W N, Yang T, Chai T Y. Adaptive optimal output regulation based on adaptive dynamic programming and gradient descent method[J]. Control and Decision, 2023,

- 38(8): 2425-2432.)
- [13] 赵振根, 程磊. 基于增量式 Q 学习的固定翼无人机跟踪控制性能优化[J]. 控制与决策, 2024, 39(2): 391-400. (Zhao Z G, Cheng L. Performance optimization for tracking control of fixed-wing UAV with incremental Q -learning[J]. Control and Decision, 2024, 39(2): 391-400.)
- [14] 温广辉, 杨涛, 周佳玲, 等. 强化学习与自适应动态规划: 从基础理论到多智能体系统中的应用进展综述[J]. 控制与决策, 2023, 38(5): 1200-1230. (Wen G H, Yang T, Zhou J L, et al. Reinforcement learning and adaptive/approximate dynamic programming: A survey from theory to applications in multi-agent systems[J]. Control and Decision, 2023, 38(5): 1200-1230.)
- [15] Jiang Y, Fan J L, Chai T Y, et al. Dual-rate operational optimal control for flotation industrial process with unknown operational model[J]. IEEE Transactions on Industrial Electronics, 2019, 66(6): 4587-4599.
- [16] Dai W, Zhang L Z, Fu J, et al. Dual-rate adaptive optimal tracking control for dense medium separation process using neural networks[J]. IEEE Transactions on Neural Networks and Learning Systems, 2021, 32(9): 4202-4216.
- [17] Li J N, Chai T Y, Lewis F L, et al. Off-policy Q -learning: Set-point design for optimizing dual-rate rougher flotation operational processes[J]. IEEE Transactions on Industrial Electronics, 2018, 65(5): 4092-4102.
- [18] Dai W, Li T Y, Zhang L Z, et al. Multi-rate layered operational optimal control for large-scale industrial processes[J]. IEEE Transactions on Industrial Informatics, 2022, 18(7): 4749-4761.
- [19] Abdelgalil M, Ochoa D E, Poveda J I. Multi-time scale control and optimization via averaging and singular perturbation theory: From ODEs to hybrid dynamical systems[J]. Annual Reviews in Control, 2023, 56: 100926.
- [20] Shen H, Peng C J, Yan H C, et al. Data-driven near optimization for fast sampling singularly perturbed systems[J]. IEEE Transactions on Automatic Control, 2024, 69(7): 4689-4694.
- [21] Kokotović P, Khalil H K, O'Reilly J. Singular perturbation methods in control: Analysis and design[M]. Philadelphia: Society for Industrial and Applied Mathematics, 1999.
- [22] Xue W Q, Fan J L, Lopez V G, et al. Off-policy reinforcement learning for tracking in continuous-time systems on two time scales[J]. IEEE Transactions on Neural Networks and Learning Systems, 2021, 32(10): 4334-4346.
- [23] He W L, Liang K, Qian F, et al. Designing distributed impulsive controller for networked singularly perturbed systems[J]. IEEE Transactions on Automatic Control, 2023, 68(6): 3680-3686.
- [24] Granzotto M, Postoyan R, Busoniu L, et al. Finite-horizon discounted optimal control: Stability and performance[J]. IEEE Transactions on Automatic Control, 2021, 66(2): 550-565.
- [25] Zhao J G, Yang C Y, Gao W N, et al. Optimal dynamic controller design for linear quadratic tracking problems[J]. IEEE Transactions on Automatic Control, 2024, 69(6): 4021-4027.
- [26] Lewis F L, Vrabie D L, Syrmos V L. Optimal control[M]. New York: John Wiley & Sons, 2012.
- [27] Kleinman D. On an iterative technique for Riccati equation computations[J]. IEEE Transactions on Automatic Control, 1968, 13(1): 114-115.
- [28] Jiang Y, Jiang Z P. Robust Adaptive Dynamic Programming[M]. New York: Wiley, 2017.

作者简介

周林娜 (1979-), 女, 副教授, 博士, 博士生导师, 主要研究方向为工业工程运行优化控制、奇异摄动系统稳定性分析与控制, E-mail: linnazhou@cumt.edu.cn;

蔡天赐 (2001-), 男, 硕士生, 主要研究方向为工业工程运行优化控制, E-mail: tiancicai@cumt.edu.cn;

赵建国 (1993-), 男, 助理研究员, 博士, 主要研究方向为多时间尺度系统、强化学习最优控制, E-mail: jianguozhao@cumt.edu.cn;

熊梦辉 (1995-), 男, 助理研究员, 博士, 主要研究方向为分布式优化、网络化控制, E-mail: xiongmenghui@cumt.edu.cn;

杨春雨 (1979-), 男, 教授, 博士, 主要研究方向为智能系统控制技术、奇异摄动系统优化控制, E-mail: chunyu yang@cumt.edu.cn.