

数据驱动的保证收敛速率最优输出调节

姜艺¹ 范家璐¹ 柴天佑¹

摘要 针对具有外部系统扰动的线性离散时间系统的输出调节问题,提出了可保证收敛速率的数据驱动最优输出调节方法,包括状态可在线测量系统的基于状态反馈的算法,与状态不可在线测量系统的基于输出反馈的算法.首先,该问题被分解为输出调节方程求解问题与反馈控制律设计问题,基于输出调节方程的解,通过引入收敛速率参数,建立了可保证收敛速率的最优控制问题,通过求解该问题得到具有保证收敛速率的输出调节器.之后,利用强化学习的方法,设计基于值迭代的数据驱动状态反馈控制器,学习得到基于状态反馈的最优输出调节器.对于状态无法在线测量的被控对象,利用历史输入输出数据对状态进行重构,并以此为基础设计基于值迭代的数据驱动输出反馈控制器.仿真结果验证了所提方法的有效性.

关键词 保证收敛速率,最优输出调节,强化学习,值迭代

引用格式 姜艺,范家璐,柴天佑.数据驱动的保证收敛速率最优输出调节.自动化学报,2022,48(4):980-991

DOI 10.16383/j.aas.c200932

Data-driven Optimal Output Regulation With Assured Convergence Rate

JIANG Yi¹ FAN Jia-Lu¹ CHAI Tian-You¹

Abstract This paper investigates the output regulation problem for linear discrete-time systems with disturbances caused by exosystem and proposes data-driven optimal output regulation approaches with assured convergence rate, including the state feedback based algorithm for the system whose state can be measured online, and the output feedback based algorithm for the system whose state cannot be measured online. Firstly, this problem is decomposed into an output regulation equation solving problem and a feedback control law design problem. Based on the solutions of the output regulation equation, by introducing the convergence rate parameter, an optimal control problem with assured convergence rate is formulated and an assured convergence rate output regulator can be obtained by solving this problem. Then, by using the reinforcement learning approach, this paper designs a value iteration based data-driven state feedback controller which can learn the state feedback based optimal output regulator. For the systems whose states cannot be measured online, the state is reconstructed by using historical input and output data, and a data-driven output feedback controller based on value iteration is designed. Simulation results show the effectiveness of the proposed approaches.

Key words Assured convergence rate, optimal output regulation, reinforcement learning, value iteration

Citation Jiang Yi, Fan Jia-Lu, Chai Tian-You. Data-driven optimal output regulation with assured convergence rate. *Acta Automatica Sinica*, 2022, 48(4): 980-991

在实际的控制器设计问题中,通常是希望将被控对象的输出跟踪给定的设定值或给定的期望轨迹,即实现输出跟踪.对于前者,PID控制器^[1]、模型预测控制器^[2]是一类经典的解决方案.对于后者,

该问题通常可以建立成一类输出调节问题^[3-6],该问题的目标通常包括两部分,设计稳定的控制器使得输出信号与给定参考轨迹的误差是渐近稳定的,并且能够完全可以克服外部系统所产生扰动信号对系统所产生的影响.然而,解决输出调节问题通常依赖于已知的精确模型参数,而在一些特殊情况下该要求是难以满足的.

针对模型未知的被控对象的输出跟踪问题,一些专家学者提出了基于自适应的控制方法,如模型参考自适应控制^[7]、无模型自适应控制^[8]、神经网络自适应控制^[9],这些方法可以在部分模型知识未知的情况下,很好的实现输出跟踪.而在有些情况下,控制器目标需要使得最小化给定的性能指标,同时希望系统的动态性能满足一定要求,这使得需要设计最优自适应控制器.

收稿日期 2020-11-23 录用日期 2021-01-27

Manuscript received November 23, 2020; accepted January 27, 2021

国家自然科学基金(61991404, 61991403, 61991400, 61533015),中央高校基本科研专项资金(N180804001),2020年度辽宁省科技重大专项计划(2020JH1/10100008)资助

Supported by National Natural Science Foundation of China (61991404, 61991403, 61991400, 61533015), the Fundamental Research Funds for the Central Universities (N180804001) and 2020 Science and Technology Major Project of Liaoning Province (2020JH1/10100008)

本文责任编辑 倪茂林

Recommended by Associate Editor NI Mao-Lin

1. 东北大学流程工业综合自动化国家重点实验室 沈阳 110819

1. State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang 110819

为解决最小化给定的性能指标问题, 一些专家学者提出了基于强化学习的自适应控制方法, 该方法通过与未知被控对象的交互来更新控制策略, 使得控制器是最优的. 对于跟踪问题, 主要有两类基于强化学习的方法, 一类是将跟踪问题定义为一类最优二次型跟踪问题, 另一类是基于输出调节理论的最优输出调节问题. 利用前一类方法, 文献 [10-11] 与文献 [12-15] 分别解决了连续与离散线性系统的最优跟踪控制问题, 文献 [16] 与文献 [17-19] 分别解决了连续与离散非线性系统的最优跟踪控制问题. 利用后一类方法, 文献 [20-23] 与文献 [24-26] 分别解决了连续与离散线性系统的最优输出调节问题, 文献 [27] 与文献 [28] 分别解决了连续与离散非线性系统的最优输出调节问题. 上述方法是基于状态反馈与策略迭代的方法, 而对于系统状态难以在线测量的系统, 上述方法不能直接应用, 针对这个问题, 文献 [29] 与文献 [30] 分别设计了基于输出反馈的控制器解决了最优跟踪控制问题与最优输出调节问题. 对于动态性能要求, 文献 [31] 针对单无人机对单目标的环航跟踪问题, 设计了飞行轨迹快速收敛到期望航迹的控制器. 文献 [32] 通过设计状态反馈和动态输出反馈控制, 研究了机器人系统的有限时间控制问题. 然而, 上述文献需要利用系统的动态模型参数来设计合适的 Lyapunov 函数.

为了使系统的动态特性满足预先给定的要求, 同时实现最优自适应控制, 本文提出保证收敛速率的数据驱动线性离散系统最优输出调节方法, 该方法不需要部分模型知识, 与文献 [24-25] 中的方法与被控对象相比, 该算法不需要稳定的初始控制律, 同时输出方程中输入到输出的前馈增益矩阵不等于 0, 利用在线的状态数据、输入数据, 或者在线的输出、输入数据求解得到基于状态反馈与输出反馈最优的输出调节器, 并保证跟踪误差的收敛速率满足预先给定的要求.

本文结构如下: 第 1 节给出离散线性系统的最优输出调节问题描述, 第 2 节与第 3 节分别进行基于状态反馈与输出反馈的自适应最优输出调节器设计, 第 4 节给出设计方法的收敛性与系统闭环稳定性分析, 第 5 节利用仿真实验验证本文设计方法的有效性, 第 6 节为结论.

符号说明: \mathbf{R} 与 \mathbf{N} 分别代表实数集与非 0 自然数集, 对于矩阵 $X, Y \in \mathbf{R}^{n \times n}$, $n \in \mathbf{N}$, $X > 0$ ($X \geq 0$) 表示 X 是正定的 (半正定的), $X > Y$ ($X \geq Y$) 表示 $X - Y$ 是正定的 (半正定的), X^{-1} 表示 X 的逆, $\sigma(X)$ 表示 X 的谱. $\|\cdot\|$ 表示矩阵或向量范数, 对于矩阵 $X \in \mathbf{R}^{m \times n}$, $m, n \in \mathbf{N}$, X^T 表示 X 的转置,

$\text{vec}(X) = [x_1^T, x_2^T, \dots, x_n^T]^T$, 其中 x_i , $i = 1, \dots, n$ 为矩阵 X 的第 i 列, \otimes 表示 Kronecker 积, 对于对称矩阵 $X \in \mathbf{R}^{n \times n}$, $\text{vecs}(X) = [x_{11}, x_{12}, \dots, x_{(n-1)n}, x_{nn}]^T \in \mathbf{R}^{(1/2)n(n+1)}$, 对于向量 $v \in \mathbf{R}^n$, $\text{vecv}(v) = [v_1^2, 2v_1v_2, \dots, 2v_1v_n, v_2^2, 2v_2v_3, \dots, 2v_{n-1}v_n, v_n^2]^T \in \mathbf{R}^{(1/2)n(n+1)}$, 其中 v_i , $i = 1, \dots, n$ 为向量 v 的第 i 个元素.

1 控制问题描述

考虑如下受扰动的线性离散系统

$$x(k+1) = Ax(k) + Bu(k) + Dw(k) \quad (1)$$

$$y(k) = Cx(k) + Su(k) \quad (2)$$

其中, $x \in \mathbf{R}^{n_x}$, $u \in \mathbf{R}^{n_u}$, $y \in \mathbf{R}^{n_y}$, $w \in \mathbf{R}^{n_w}$ 分别为系统的状态, 控制输入, 输出, 外部系统状态. $A \in \mathbf{R}^{n_x \times n_x}$, $B \in \mathbf{R}^{n_x \times n_u}$, $D \in \mathbf{R}^{n_x \times n_w}$, $C \in \mathbf{R}^{n_y \times n_x}$, $S \in \mathbf{R}^{n_y \times n_u}$ 为常数矩阵. 外部系统动态及其所产生的设定值为:

$$w(k+1) = Ew(k) \quad (3)$$

$$y_d(k) = -Fw(k) \quad (4)$$

其中, $E \in \mathbf{R}^{n_w \times n_w}$ 为常数矩阵, 且其特征值都在单位圆上. $y_d \in \mathbf{R}^{n_y}$ 为参考信号, $F \in \mathbf{R}^{n_y \times n_w}$ 为常数矩阵. 基于此, 跟踪误差可以表示为:

$$e(k) = y(k) - y_d(k) = Cx(k) + Su(k) + Fw(k) \quad (5)$$

针对此系统, 有如下假设:

假设 1. (A, B) 是可控的, (A, C) 是可观测的.

假设 2. $\text{rank} \left(\begin{bmatrix} A - \lambda I & B \\ C & S \end{bmatrix} \right) = n_x + n_y$, $\forall \lambda \in \sigma(E)$.

假设 3. 矩阵 E 的特征值都在单位圆上且互不重复.

假设 4. $\left(\begin{bmatrix} A & D \\ 0 & E \end{bmatrix}, [C \ F] \right)$ 是可观测的.

传统的输出调节问题的控制器设计目标为使得跟踪误差 $e(k)$ 是渐近稳定的, 即 $\lim_{k \rightarrow \infty} e(k) = 0$. 本文目标为利用外部系统数据 $w(k)$, 系统输入 $u(k)$, 系统状态 $x(k)$ 或系统输出 $y(k)$ 设计最优输出调节器, 使得跟踪误差 $e(k)$ 是渐近稳定的, 同时期望跟踪误差 $e(k)$ 的收敛速率快于 γ^{-k} , 其中 $\gamma > 1$. 该问题可以定义为求解如下问题.

问题 1. 针对被控对象 (1) ~ (2), 对应的外部系统为 (3) ~ (4), 设计控制器 $u(k)$ 使得跟踪误差满足

$$\lim_{k \rightarrow \infty} \gamma^k e(k) = 0 \quad (6)$$

为解决该问题, 根据输出调节理论^[3, 33], 该问题的输出调节方程为

$$XE = AX + BU + D \quad (7)$$

$$0 = CX + SU + F \quad (8)$$

其中, $X \in \mathbf{R}^{n_x \times n_w}$ 与 $U \in \mathbf{R}^{n_u \times n_w}$ 为输出调节方程的待求解未知数. 利用 Kronecker 积, 输出调节方程 (7) ~ (8) 可写为

$$\Gamma \eta = \vartheta \quad (9)$$

其中

$$\Gamma = E^T \otimes \begin{bmatrix} I_{n_x} & 0_{n_x \times n_u} \\ 0_{n_y \times n_x} & 0_{n_y \times n_u} \end{bmatrix} - I_{n_w} \otimes \begin{bmatrix} A & B \\ C & S \end{bmatrix}$$

$$\eta = \text{vec} \left(\begin{bmatrix} X \\ U \end{bmatrix} \right), \vartheta = \text{vec} \left(\begin{bmatrix} D \\ F \end{bmatrix} \right)$$

基于假设 2 可知, Γ 是行满秩的, 输出调节方程 (7) ~ (8) 是有解的^[33]. 基于该解, 并同时考虑控制器设计要求为使得跟踪误差 $e(k)$ 的收敛速率快于 γ^{-k} , 定义新系统为

$$\bar{x}(k+1) = \bar{A}\bar{x}(k) + \bar{B}\bar{u}(k) \quad (10)$$

$$\bar{e}(k) = C\bar{x}(k) + S\bar{u}(k) \quad (11)$$

其中, $\bar{x}(k) = \gamma^k(x(k) - Xw(k))$, $\bar{A} = \gamma A$, $\bar{B} = \gamma B$, $\bar{u}(k) = \gamma^k(u(k) - Uw(k))$, $\bar{e}(k) = \gamma^k e(k)$.

基于新系统 (10) ~ (11), 建立如下最优控制问题与约束最优化问题. 通过求解该问题, 可以保证式 (6) 成立, 即跟踪误差 $e(k)$ 的收敛速率快于 γ^{-k} , 该性质将会在闭环系统分析部分进行证明.

问题 2^[34]. 针对系统 (10) ~ (11), 给定 $Q \geq 0$, $R \geq 0$, 设计基于状态反馈与输出反馈的最优控制输入 $\bar{u}(k)$, 使得如下性能指标最小

$$\begin{aligned} \min_{\bar{u}} V(k) &= \sum_{i=k}^{\infty} (\bar{e}^T(i)Q\bar{e}(i) + \bar{u}^T(i)R\bar{u}(i)) \\ \text{s.t.} \quad &\begin{cases} \bar{x}(k+1) = \bar{A}\bar{x}(k) + \bar{B}\bar{u}(k) \\ \bar{e}(k) = C\bar{x}(k) + S\bar{u}(k) \end{cases} \end{aligned} \quad (12)$$

问题 3. 给定 $M > 0$, 寻找出一组输出调节方程 (7)~(8) 的解 X 和 U 使得如下性能指标最小

$$\begin{cases} \min_{\eta} J = \eta^T M \eta \\ \text{s.t.} \quad \Gamma \eta = \vartheta \end{cases} \quad (13)$$

注 1. 在问题 2 中, 选择 Q 与 R 使得如下的广义特征值问题的解不在单位圆上^[35]

$$\lambda \begin{bmatrix} I & 0 & 0 \\ 0 & \bar{A}^T & 0 \\ 0 & -\bar{B}^T & 0 \end{bmatrix} - \begin{bmatrix} \bar{A} & 0 & \bar{B} \\ -C^T Q C & I & -C^T Q S \\ S^T Q C & 0 & R \end{bmatrix}$$

2 基于状态反馈的自适应最优输出调节器设计

本节在被控对象状态方程 (1) 中矩阵 A 、 B 、 D 、 E 未知、被控对象输出方程 (2) 中矩阵 C 、 S 与 F 已知的情况下, 设计数据驱动的基于状态反馈的最优自适应输出调节器. 首先给出基于状态反馈的最优输出调节器的解, 之后利用该解的求解形式, 设计数据驱动的基于值迭代的自适应最优输出调节器. 值得注意的是, 由于本节所设计的是基于状态反馈的最优输出调节器, 因此需要利用状态计算跟踪误差, 故矩阵 C 、 S 与 F 已知的假设是合理的.

2.1 基于状态反馈与模型的最优输出调节器

本小节首先求解输出调节方程 (7) ~ (8), 引入两个 Sylvester 映射 $\Omega: \mathbf{R}^{n_x \times n_w} \rightarrow \mathbf{R}^{n_x \times n_w}$, $\bar{\Omega}: \mathbf{R}^{n_x \times n_w} \times \mathbf{R}^{n_u \times n_w} \rightarrow \mathbf{R}^{n_x \times n_w}$, 为

$$\Omega(X) = XE - AX \quad (14)$$

$$\bar{\Omega}(X, U) = XE - AX - BU \quad (15)$$

基于 Sylvester 映射, 可以给出方程 (8) 的通解形式. 选择两个矩阵序列 $X_i \in \mathbf{R}^{n_x \times n_w}$ 与 $U_i \in \mathbf{R}^{n_u \times n_w}$, 其中 $i = 0, 1, \dots, m+1$, m 为 $I_{n_w} \otimes [C, S]$ 的零空间的维数. $X_0 = 0_{n_x \times n_w}$, $U_0 = 0_{n_u \times n_w}$, X_1 与 U_1 满足 $-F = CX_1 + SU_1$, 当 $i = 2, 3, \dots, m+1$ 时, $\text{vec} \left([X_i^T, U_i^T]^T \right)$ 为 $I_{n_w} \otimes [C, S]$ 的基底, 即 $CX_i + SU_i = 0$. 则方程 (8) 的通解为

$$(X, U) = (X_1, U_1) + \sum_{i=2}^{m+1} \alpha_i (X_i, U_i) \quad (16)$$

其中, $\alpha_i \in \mathbf{R}$. 由 $\bar{\Omega}(X, U)$ 的定义与式 (7) 可知

$$\bar{\Omega}(X, U) = D, \quad \bar{\Omega}(\alpha_i X, \alpha_i U) = \alpha_i \bar{\Omega}(X, U)$$

$$\bar{\Omega}(X_i + X_j, U_i + U_j) =$$

$$(X_i + X_j)E - A(X_i + X_j) - B(U_i + U_j) =$$

$$\bar{\Omega}(X_i, U_i) + \bar{\Omega}(X_j, U_j)$$

基于式 (16) 可将式 (15) 写为

$$\bar{\Omega}(X, U) = \bar{\Omega}(X_1, U_1) + \sum_{i=2}^{m+1} \alpha_i \bar{\Omega}(X_i, U_i) = D \quad (17)$$

至此, 输出调节方程 (7) ~ (8) 可以写为

$$\Lambda \chi = \xi \quad (18)$$

其中

$$\Lambda = \begin{bmatrix} \text{vec}(\bar{\Omega}(X_2, U_2)) & \cdots \\ \text{vec}([X_2^T, U_2^T]^T) & \cdots \\ \text{vec}(\bar{\Omega}(X_{m+1}, U_{m+1})) & 0 \\ \text{vec}([X_{m+1}^T, U_{m+1}^T]^T) & -I_{n_w(n_x+n_u)} \end{bmatrix}$$

$$\chi = [\alpha_2 \quad \cdots \quad \alpha_{m+1} \quad \eta^T]^T$$

$$\xi = \begin{bmatrix} \text{vec}(-\bar{\Omega}(X_1, U_1) + D) \\ -\text{vec}([X_1^T, U_1^T]^T) \end{bmatrix} = \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix}$$

利用矩阵行变换, 可以将 (18) 重写为

$$\begin{bmatrix} \bar{\Lambda}_{11} & \bar{\Lambda}_{12} \\ \bar{\Lambda}_{21} & \bar{\Lambda}_{22} \end{bmatrix} \chi = \begin{bmatrix} \bar{\xi}_1 \\ \bar{\xi}_2 \end{bmatrix} \quad (19)$$

其中, $\bar{\Lambda}_{21} \in \mathbf{R}^{m \times m}$ 为非奇异矩阵, 则上式可以通过如下方程求解

$$\Pi \eta = \Psi \quad (20)$$

其中, $\Pi = -\bar{\Lambda}_{11}\bar{\Lambda}_{21}^{-1}\bar{\Lambda}_{22} + \bar{\Lambda}_{12}$, $\Psi = -\bar{\Lambda}_{11}\bar{\Lambda}_{21}^{-1}\bar{\xi}_2 + \bar{\xi}_1$. 利用拉格朗日乘子法, 可以将问题 3 中的约束最优优化问题 (13) 转化为

$$\min_{\eta} J = \eta^T M \eta + \lambda^T (\Pi \eta - \Psi) \quad (21)$$

对上述性能指标 J 求对于 η 与 λ^T 偏导, 可得

$$\frac{\partial J}{\partial \eta} = 2M\eta + \Pi^T \lambda \quad (22)$$

$$\frac{\partial J}{\partial \lambda^T} = \Pi \eta - \Psi \quad (23)$$

令式 (22) 和式 (23) 等于 0, 可得

$$\begin{bmatrix} 2M & \Pi^T \\ \Pi & 0 \end{bmatrix} \begin{bmatrix} \eta \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ \Psi \end{bmatrix} \quad (24)$$

利用该式可以得到输出调节方程 (7) ~ (8) 的解 X 和 U .

以上为基于模型的输出调节方程 (7) ~ (8) 的求解, 与文献 [32] 直接求解输出调节方程不同, 式 (20) 中的求解方法将会为第 2.2 节中自适应控制器设计提供指导.

基于输出调节方程 (7) ~ (8) 的解 X 和 U , 则基于状态反馈的最优输出调节问题可以总结为问题 2. 对于问题 2, 该问题为标准的最优控制问题. 基于假设 1, 可知 $\text{rank}[A - \lambda I, B] = n_x, \forall \lambda$, 那么显然有 $\text{rank}[\bar{A} - \lambda I, \bar{B}] = n_x, \forall \lambda$, 即 (\bar{A}, \bar{B}) 为可控的. 因此, 基于最优控制原理^[36], 假设存在一个可控的矩阵 K 与控制输入 $\bar{u}(k) = -K\bar{x}(k)$ 使得

$$V(k) = \bar{e}^T(k)Q\bar{e}(k) + \bar{u}^T(k)R\bar{u}(k) + V(k+1) = \bar{x}^T(k)P\bar{x}(k) \quad (25)$$

其中, $P \geq 0$. 上述贝尔曼方程可以写为

$$P = (C - SK)^T Q (C - SK) + K^T R K + (\bar{A} - \bar{B}K)^T P (\bar{A} - \bar{B}K) \quad (26)$$

通过使得 $\partial P / \partial K = 0$ 可得最优反馈增益为

$$K^* = (R + \bar{B}^T P^* \bar{B} + S^T Q S)^{-1} (\bar{B}^T P^* \bar{A} + S^T Q C) \quad (27)$$

其中, P^* 为如下 Riccati 方程的解

$$P = C^T Q C + \bar{A}^T P \bar{A} - (\bar{A}^T P \bar{B} + C^T Q S) (R + \bar{B}^T P \bar{B} + S^T Q S)^{-1} (\bar{B}^T P \bar{A} + S^T Q C) \quad (28)$$

对应的最优控制输入为:

$$u(k) = -K^* x(k) + (U + K^* X) w(k) := -K^* x(k) + L^* w(k) \quad (29)$$

然而, 直接求解 Riccati 方程比较复杂, 针对此问题, 该小节利用基于值迭代的算法求解, 其收敛性性质见如下引理.

算法 1. 基于模型的值迭代状态反馈最优输出调节算法

初始化: 选择任意的初始控制律 K_0 , 终止条件常数 $\varepsilon > 0$, 矩阵序列 $X_i \in \mathbf{R}^{n_x \times n_w}$ 与 $U_i \in \mathbf{R}^{n_u \times n_w}$, 半正定矩阵 $P_0, j \leftarrow 0$;

最优反馈增益计算: 利用如下迭代算法计算最优反馈增益;

1) 计算 P_{j+1} ,

$$P_{j+1} = (C - SK_j)^T Q (C - SK_j) + K_j^T R K_j + (\bar{A} - \bar{B}K_j)^T P_j (\bar{A} - \bar{B}K_j) \quad (30)$$

2) 计算 K_{j+1} ,

$$K_{j+1} = (R + \bar{B}^T P_{j+1} \bar{B} + S^T Q S)^{-1} (\bar{B}^T P_{j+1} \bar{A} + S^T Q C) \quad (31)$$

3) 判断 $\|P_{j+1} - P_j\| < \varepsilon$ 是否成立, 如果成立则停止迭代, 反之则继续重复计算上述两步, 并令 $j \leftarrow j + 1$;

最优前馈增益计算: 利用式 (9) 或式 (24) 求解输出调节方程 (7) ~ (8) 的解 X 和 U , 进而得到最优前馈增益 L^* .

引理 1. 在假设 1 成立的条件下, 通过算法 1 的式 (30) ~ (31) 计算得到的序列 $\{P_j\}_{j=0}^{\infty}$ 与 $\{K_j\}_{j=0}^{\infty}$ 最终会收敛至其最优值, 即 $\lim_{j \rightarrow \infty} P_j = P^*$, $\lim_{j \rightarrow \infty} K_j = K^*$.

证明. 文献 [37] 给出了当 $S = 0$ 时的收敛性证明, 本文将简述 $S \neq 0$ 时的收敛性证明. 首先将式 (28) 与式 (26) 定义为

$$P = g(P) \tag{32}$$

$$P = L(K, P) \tag{33}$$

同时定义

$$M(K, P) = (\bar{A} - \bar{B}K)^T P (\bar{A} - \bar{B}K) \tag{34}$$

由于在式 (26) 中 P 是 K 的二次型, 可得

$$g(P) = \min_K L(K, P) = L(K_P, P) \leq L(K, P)$$

其中

$$K_P = (R + \bar{B}^T P \bar{B} + S^T Q S)^{-1} (\bar{B}^T P \bar{A} + S^T Q C)$$

根据上式可知, 对于任意的 $X \leq Y$, 有

$$g(X) = L(K_X, X) \leq L(K_Y, Y) = g(Y)$$

考虑序列 $\{Q_j\}_{j=0}^\infty$, 其中 $Q_0 = 0$, 可得

$$Q_{j+1} = g(Q_j) \leq g(Q_{j+1}) = Q_{j+2} \tag{35}$$

$$Q_{j+1} \leq M(K^*, Q_j) + (K^*)^T R K^* + (C - S K^*)^T Q (C - S K^*) \tag{36}$$

根据上式可知, 由于 $\bar{A} - \bar{B}K^*$ 的特征值都在单位圆内, 序列 $\{Q_j\}_{j=0}^\infty$ 是单调递增且存在上界, 即 $\lim_{j \rightarrow \infty} Q_j = P^*$. 之后考虑序列 $\{R_j\}_{j=0}^\infty$, 其中 $R_0 \geq P^*$ 且 $R_0 \geq P_0$, 可得

$$R_{j+1} = g(R_j) \geq g(P^*) = P^* \tag{37}$$

$$R_{j+1} - P^* \leq M(K^*, R_j - P^*) \tag{38}$$

同理可知, 序列 $\{R_j\}_{j=0}^\infty$ 是单调递减且存在下界, 即 $\lim_{j \rightarrow \infty} R_j = P^*$. 综上所述, 可得

$$P^* = \lim_{j \rightarrow \infty} Q_j \leq \lim_{j \rightarrow \infty} P_j \leq \lim_{j \rightarrow \infty} R_j = P^* \tag{39}$$

根据夹逼定理, 可得 $\lim_{j \rightarrow \infty} P_j = P^*$, 进而可得 $\lim_{j \rightarrow \infty} K_j = K^*$. \square

注 2. 在传统的基于输出调节原理的输出调节方法中, $\gamma = 1$, 对应的 Riccati 方程 (28) 可解条件为 (A, B) 是可镇定的. 当 $\gamma > 1$ 且 (A, B) 是可镇定时, 选择 $\gamma < \bar{\gamma}$, 其中 $1/\bar{\gamma}$ 大于 A 的最大不可控稳定特征值, 可以保证 (\bar{A}, \bar{B}) 是可镇定的.

注 3. 对于基于策略迭代的算法^[24-25, 38], 其初始控制律 K_0 要求矩阵 $\bar{A} - \bar{B}K_0$ 是稳定的, 即 $A - BK_0$ 的特征值在以原点为圆心, 半径为 $1/\gamma$ 的圆内, 当矩阵 A, B 已知时, 选择满足该条件的初始控制律 K_0 是很容易的, 然而, 当矩阵 A, B 未知时, 初始控制律的选择则更加严格. 因此, 本文使用基于值迭代的算法, 该算法的初始控制律 K_0 可以是任意的,

同时该算法不用重复求解 Lyapunov 函数^[24-25, 38].

以上为基于模型的问题求解方法, 该求解方法将会为下一节中自适应控制器设计提供指导.

2.2 基于状态反馈与强化学习的自适应最优输出调节器

本小节利用第 2.1 节给出的最优输出调节器的求解形式, 设计利用在线数据的基于状态反馈与强化学习的自适应最优输出调节器, 首先定义新状态 $\bar{x}_i(k) = \hat{x}_i(k) - X_i \hat{w}(k)$, 其中 $\hat{x}(k) = \gamma^k x(k)$, $\hat{w}(k) = \gamma^k w(k)$, 基于该状态, 可得

$$\begin{aligned} \bar{x}_i(k+1) &= \bar{A} \bar{x}_i(k) + \bar{B} \hat{u}(k) + \\ &\quad \gamma (D - \Omega(X_i)) \hat{w}(k) \end{aligned} \tag{40}$$

基于以上动态方程, 可得

$$\begin{aligned} \bar{x}_i^T(k+1) P_j \bar{x}_i(k+1) &= \\ \bar{x}_i^T(k) \bar{A}^T P_j \bar{A} \bar{x}_i(k) &+ \hat{u}^T(k) \bar{B}^T P_j \bar{B} \hat{u}(k) + \\ \gamma^2 \hat{w}^T(k) (D - \Omega(X_i))^T P_j (D - \Omega(X_i)) \hat{w}(k) &+ \\ 2 \bar{x}_i^T(k) \bar{A}^T P_j \bar{B} \hat{u}(k) &+ \\ 2 \gamma \hat{u}^T(k) \bar{B}^T P_j (D - \Omega(X_i)) \hat{w}(k) &+ \\ 2 \gamma \bar{x}_i^T(k) \bar{A}^T P_j (D - \Omega(X_i)) \hat{w}(k) & \end{aligned} \tag{41}$$

通过定义

$$\begin{aligned} L_{1j} &= \bar{A}^T P_j \bar{A}, L_{2j} = \bar{B}^T P_j \bar{B}, L_{3j} = \bar{A}^T P_j \bar{B} \\ L_{4ij} &= \bar{A}^T P_j (D - \Omega(X_i)), L_{5ij} = \bar{B}^T P_j (D - \Omega(X_i)) \\ L_{6ij} &= (D - \Omega(X_i))^T P_j (D - \Omega(X_i)) \end{aligned}$$

$$\varphi_j^i(k) = \begin{bmatrix} \gamma^{-2k} \bar{x}_i^T(k+1) P_j \bar{x}_i(k+1) \\ \gamma^{-2k-2} \bar{x}_i^T(k+2) P_j \bar{x}_i(k+2) \\ \vdots \\ \gamma^{-2k-2s} \bar{x}_i^T(k+1+s) P_j \bar{x}_i(k+1+s) \end{bmatrix}$$

$$\psi_j^i(k) = \begin{bmatrix} \Phi_{01} & \Phi_{02} & \cdots & \Phi_{06} \\ \Phi_{11} & \Phi_{12} & \cdots & \Phi_{16} \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_{s1} & \Phi_{s2} & \cdots & \Phi_{s6} \end{bmatrix}$$

其中

$$\begin{aligned} \Phi_{l1} &= \gamma^{-2k-2l} \text{vecv}(\bar{x}_i(k+l)) \\ \Phi_{l2} &= \text{vecv}(u(k+l)) \\ \Phi_{l3} &= 2\gamma^{-k-l} u^T(k+l) \otimes \bar{x}_i^T(k+l) \\ \Phi_{l4} &= 2\gamma^{-k-l+1} w^T(k+l) \otimes \bar{x}_i^T(k+l) \\ \Phi_{l5} &= 2\gamma w^T(k+l) \otimes u^T(k+l) \\ \Phi_{l6} &= \gamma^2 \text{vecv}(w(k+l)) \end{aligned}$$

可将式 (41) 转化为如下方程组

$$\begin{aligned} \psi_j^i(k) [\text{vecs}(L_{1j}); \text{vecs}(L_{2j}); \text{vecs}(L_{3j}); \\ \text{vecs}(L_{4ij}); \text{vecs}(L_{5ij}); \text{vecs}(L_{6ij})] = \varphi_j^i(k) \quad (42) \end{aligned}$$

当在线数据满足一定要求时, 上述方程组可由最小二乘方法求解. 如下引理给出了方程组 (42) 具有唯一解的条件.

引理 2. 方程组 (42) 可解并具有唯一解, 当且仅当

$$\text{rank}(\psi_j^i(k)) = \frac{1}{2}(n_x + n_u + n_w)(n_x + n_u + n_w + 1)$$

当引理 2 成立时, 方程组可以由下式求解, 为

$$\begin{aligned} [\text{vecs}(L_{1j}); \text{vecs}(L_{2j}); \text{vecs}(L_{3j}); \text{vecs}(L_{4ij}); \\ \text{vecs}(L_{5ij}); \text{vecs}(L_{6ij})] = \\ (\psi_j^{iT}(k)\psi_j^i(k))^{-1}\psi_j^{iT}(k)\varphi_j^i(k) \quad (43) \end{aligned}$$

同时考虑式 (30) 与式 (31) 可得

$$\begin{aligned} P_{j+1} = L_{1j} - (L_{3j} + C^T QS)(R + L_{2j} + S^T QS)^{-1} \cdot \\ (L_{3j}^T + S^T QC) + C^T QC \quad (44) \end{aligned}$$

计算得到 P_{j+1} 后, 将其代入 $\varphi_j^i(k)$ 更新得到 $\varphi_{j+1}^i(k)$, 继而可以更新方程 (43), 重复以上步骤可以得到序列 $\{P_j\}_{j=0}^{\infty}$ 直至收敛, 对应的序列 $\{K_j\}_{j=0}^{\infty}$ 为

$$K_j = (R + L_{2j} + S^T QS)^{-1}(L_{3j}^T + S^T QC) \quad (45)$$

以上为反馈控制增益 K_j 的在线计算过程, 这部分将介绍如何在线求解输出调节方程 (7) ~ (8) 的解 X 和 U , 基于式 (17) 可得

$$\begin{aligned} \bar{A}^T P_j \bar{\Omega}(X_1, U_1) + \sum_{i=2}^{m+1} \alpha_i \bar{A}^T P_j \bar{\Omega}(X_i, U_i) = \\ \bar{A}^T P_j \Omega(X_1) - \gamma^{-1} \bar{A}^T P_j \bar{B} U_1 + \\ \sum_{i=2}^{m+1} \alpha_i \bar{A}^T P_j \Omega(X_i) - \sum_{i=2}^{m+1} \alpha_i \gamma^{-1} \bar{A}^T P_j \bar{B} U_i = \\ L_{40j} - L_{41j} - \gamma^{-1} L_{3j} U_1 + \\ \sum_{i=2}^{m+1} \alpha_i (L_{40j} - L_{4ij} - \gamma^{-1} L_{3j} U_i) = \\ \bar{A}^T P_j D = L_{40j} \quad (46) \end{aligned}$$

利用上式, 可将输出调节方程 (7) ~ (8) 写为

$$\Lambda_j \chi = \xi_j \quad (47)$$

其中

$$\begin{aligned} \Lambda_j = \begin{bmatrix} \text{vec}(L_{40j} - L_{42j} - \gamma^{-1} L_{3j} U_2) & \cdots \\ \text{vec}([X_2^T, U_2^T]^T) & \cdots \\ \text{vec}(L_{40j} - L_{4(m+1)j} - \gamma^{-1} L_{3j} U_{m+1}) & 0 \\ \text{vec}([X_{m+1}^T, U_{m+1}^T]^T) & -I \end{bmatrix} \\ \xi_j = \begin{bmatrix} \text{vec}(L_{41j} + \gamma^{-1} L_{3j} U_1) \\ -\text{vec}([X_1^T, U_1^T]^T) \end{bmatrix} \end{aligned}$$

利用矩阵行变换, 可以将式 (47) 重写为类似式 (19) 的形式, 进而可以利用式 (20) 进行求解得到输出调节方程 (7) ~ (8) 的解 X 和 U , 最后利用式 (29) 得到前馈增益. 至此, 基于状态反馈与强化学习的自适应最优输出调节算法如下.

算法 2. 基于状态反馈与强化学习的自适应最优输出调节算法

初始化: 选择任意的初始控制律 K_0 , 终止条件常数 $\varepsilon > 0$, 半正定矩阵 P_0 , 矩阵序列 $X_i \in \mathbf{R}^{n_x \times n_w}$ 与 $U_i \in \mathbf{R}^{n_u \times n_w}$, $j \leftarrow 0$, $i \leftarrow 0$;

最优反馈控制律在线计算: 利用如下迭代算法计算最优反馈增益, 在区间 $[k, k+s]$ 利用控制输入为 $u(k) = -K_0 x(k) + n(k)$, 其中 $n(k)$ 为控制输入中添加的探测噪声, s 为使得引理 2 满足的数;

- 1) 利用式 (43) 计算得到 L_{1j} , L_{2j} , L_{3j} , L_{4ij} , L_{5ij} , L_{6ij} ;
- 2) 利用式 (44) 计算 P_{j+1} ;
- 3) 判断 $\|P_{j+1} - P_j\| < \varepsilon$ 是否成立, 如果成立则停止迭代, 并利用式 (45) 计算得到 K_j , 反之重复上述两步, 并令 $j \leftarrow j+1$;

前馈增益在线计算: 令 $i \leftarrow i+1$, 重复计算得到所有 L_{4ij} 直到 $i = m+1$, 进而利用式 (24) 进行求解得到输出调节方程 (7) ~ (8) 的解 X 和 U , 最后利用式 (29) 得到前馈增益.

注 4. 值得注意的是, $(\psi_j^{iT}(k)\psi_j^i(k))^{-1}\psi_j^{iT}(k)$ 中仅含有过程数据, 因此, 该值在迭代过程中对于固定 i 仅需要计算 1 次, 相较于基于策略迭代的方法, 本文方法虽然迭代步数多, 但每一步所需要的计算量却小一些.

注 5. 对于序列 $\{K_j\}_{j=0}^{\infty}$, 由于 $\{K_j\}_{j=0}^{\infty}$ 并不参与过程迭代, K_j 仅需要在 P_j 收敛后计算 1 次. 因此, 在该算法过程中 $u(k)$ 并不需要在线更新, 因此该方法是一类离线策略, 相较于在线策略, 该方法可以保证计算结果是无偏的^[39-40].

注 6. 探测噪声 $n(k)$ 的加入是为了使得引理 2 的条件满足, 达到充分激励的效果. 通常选择为白噪声或者正弦函数等.

3 基于输出反馈的自适应最优输出调节器设计

本节在被控对象 (1) ~ (2) 中矩阵 A, B, D, S, E, C 与 F 未知, 在 $1U$ 已知的情况下, 设计基于输出反馈的最优自适应输出调节器, 首先利用历史的输入输出数据设计重构状态^[29-30, 41], 之后设计基于值迭代的输出反馈自适应最优输出调节器.

3.1 状态重构

定义

$$\hat{A} = \begin{bmatrix} \bar{A} & \gamma D \\ 0 & \gamma E \end{bmatrix}, \quad \hat{B} = \begin{bmatrix} \bar{B} \\ 0 \end{bmatrix}$$

$$\hat{C} = [C \quad F], \quad z(k) = \begin{bmatrix} \hat{x}(k) \\ \hat{w}(k) \end{bmatrix}$$

可得

$$z(k+1) = \hat{A}z(k) + \hat{B}\hat{u}(k) \quad (48)$$

$$\bar{e}(k) = \hat{C}z(k) + S\hat{u}(k) \quad (49)$$

利用上式, 可得

$$z(k) = \hat{A}^{n_x+n_w} z(k-n_x-n_w) + \begin{bmatrix} \hat{B} & \hat{A}\hat{B} & \dots & \hat{A}^{n_x+n_w-1}\hat{B} \end{bmatrix} \cdot \begin{bmatrix} \hat{u}(k-1) \\ \hat{u}(k-2) \\ \vdots \\ \hat{u}(k-n_x-n_w) \end{bmatrix} := \hat{A}^{n_x+n_w} z(k-n_x-n_w) + U_u \hat{u}(k) \quad (50)$$

$$\bar{e}(k) = \hat{C}\hat{A}^{n_x+n_w} z(k-n_x-n_w) + \hat{C}U_u \hat{u}(k) + S\hat{u}(k) \quad (51)$$

基于上式, 考虑 $[k-1, k-n_x-n_w]$ 的输出 $\bar{e}(k)$, 可得

$$\hat{e}(k) = \begin{bmatrix} \bar{e}(k-1) \\ \bar{e}(k-2) \\ \vdots \\ \bar{e}(k-n_x-n_w) \end{bmatrix} = \begin{bmatrix} \hat{C}\hat{A}^{n_x+n_w-1} \\ \vdots \\ \hat{C}\hat{A} \\ \hat{C} \end{bmatrix} z(k-n_x-n_w) +$$

$$\begin{bmatrix} S & \hat{C}\hat{B} & \hat{C}\hat{A}\hat{B} & \dots & \hat{C}\hat{A}^{n_x+n_w-2}\hat{B} \\ 0 & S & \hat{C}\hat{B} & \dots & \hat{C}\hat{A}^{n_x+n_w-3}\hat{B} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & S \end{bmatrix} \hat{u}(k) := \bar{M}_x z(k-n_x-n_w) + \bar{M}_u \hat{u}(k) \quad (52)$$

由假设 4 可知, $\text{rank}(\bar{M}_x) = n_x + n_w$, 则有 $\bar{M}_x^T \bar{M}_x$ 是可逆的, 通过定义 $\bar{M}_x^+ = (\bar{M}_x^T \bar{M}_x)^{-1} \bar{M}_x^T$, 可得

$$z(k) = \hat{A}^{n_x+n_w} z(k-n_x-n_w) + U_u \hat{u}(k) = \hat{A}^{n_x+n_w} \bar{M}_x^+ \hat{e}(k) + (U_u - \hat{A}^{n_x+n_w} \bar{M}_x^+ \bar{M}_u) \hat{u}(k) \quad (53)$$

定义

$$\bar{M} = [I-X] \begin{bmatrix} \hat{A}^{n_x+n_w} \bar{M}_x^+ & U_u - \hat{A}^{n_x+n_w} \bar{M}_x^+ \bar{M}_u \end{bmatrix} := \begin{bmatrix} \bar{M}_1 \\ \bar{M}_2 \end{bmatrix}$$

$$\bar{z}(k) = \begin{bmatrix} \hat{e}(k) \\ \hat{u}(k) \end{bmatrix}$$

可得

$$z(k) = \begin{bmatrix} \bar{M}_1 \\ \bar{M}_2 \end{bmatrix} \bar{z}(k) \quad (54)$$

$$\bar{x}(k) = (\bar{M}_1 - X \bar{M}_2) \bar{z}(k) := \bar{M} \bar{z}(k) \quad (55)$$

3.2 基于输出反馈与强化学习的自适应最优输出调节器

由式 (29) 可知, 最优输出调节问题可由如下控制输入求解

$$\hat{u}(k) = -K^* \bar{x}(k) + U \hat{w}(k) = -(K^* \bar{M} - U \bar{M}_2) \bar{z}(k) := -\bar{K}^* \bar{z}(k) \quad (56)$$

式 (28) Riccati 方程变为

$$\bar{M}^T P \bar{M} = \bar{M}^T C^T Q C \bar{M} + \bar{M}^T \bar{A}^T P \bar{A} \bar{M} - \bar{M}^T (\bar{A}^T P \bar{B} + C^T Q S) \cdot (R + \bar{B}^T P \bar{B} + S^T Q S)^{-1} \cdot (\bar{B}^T P \bar{A} + S^T Q C) \bar{M} \quad (57)$$

上式的 Riccati 方程难以直接求解, 基于式 (55) 与动态方程

$$\bar{x}(k+1) = \bar{A} \bar{x}(k) + \bar{B} \hat{u}(k) - \bar{B} U \hat{w}(k) \quad (58)$$

可得

$$\begin{aligned}
& \bar{x}^T(k+1)P_j\bar{x}(k+1) + \bar{e}^T(k)Q\bar{e}(k) + \\
& (\hat{u}(k) - U\hat{w}(k))^T R(\hat{u}(k) - U\hat{w}(k)) = \\
& \bar{z}^T(k)(\bar{M}^T \bar{A}^T P_j \bar{A} \bar{M} + \bar{M}^T C^T Q C \bar{M} - \\
& \bar{M}^T \bar{A}^T P_j \bar{B} U \bar{M}_2 - \bar{M}_2^T U^T \bar{B}^T P_j \bar{A} \bar{M} - \\
& \bar{M}^T C^T Q S U \bar{M}_2 - \bar{M}_2^T U^T S^T Q C \bar{M} + \\
& \bar{M}_2^T U^T (\bar{B}^T P_j \bar{B} + R + S^T Q S) U \bar{M}_2) \bar{z}(k) + \\
& \hat{u}^T(k)(\bar{B}^T P_j \bar{B} + R + S^T Q S) \hat{u}(k) + \\
& 2\bar{z}^T(k)(\bar{M}^T \bar{A}^T P_j \bar{B} + \bar{M}^T C^T Q S - \\
& \bar{M}_2^T U^T (\bar{B}^T P_j \bar{B} + R + S^T Q S)) \hat{u}(k) \quad (59)
\end{aligned}$$

通过定义

$$\begin{aligned}
\bar{L}_{1j} &= \bar{M}^T \bar{A}^T P_j \bar{A} \bar{M} + \bar{M}^T C^T Q C \bar{M} - \\
& \bar{M}^T \bar{A}^T P_j \bar{B} U \bar{M}_2 - \bar{M}_2^T U^T \bar{B}^T P_j \bar{A} \bar{M} - \\
& \bar{M}^T C^T Q S U \bar{M}_2 - \bar{M}_2^T U^T S^T Q C \bar{M} + \\
& \bar{M}_2^T U^T (\bar{B}^T P_j \bar{B} + R + S^T Q S) U \bar{M}_2 \\
\bar{L}_{2j} &= \bar{B}^T P_j \bar{B} + R + S^T Q S \\
\bar{L}_{3j} &= \bar{M}^T \bar{A}^T P_j \bar{B} + \bar{M}^T C^T Q S - \\
& \bar{M}_2^T U^T (\bar{B}^T P_j \bar{B} + R + S^T Q S) \\
\bar{\varphi}_j(k) &= [f(k), f(k+1), \dots, f(k+s)]^T \\
\bar{\psi}_j(k) &= \begin{bmatrix} \bar{\Phi}_{01} & \bar{\Phi}_{02} & \bar{\Phi}_{03} \\ \bar{\Phi}_{11} & \bar{\Phi}_{12} & \bar{\Phi}_{13} \\ \vdots & \vdots & \vdots \\ \bar{\Phi}_{s1} & \bar{\Phi}_{s2} & \bar{\Phi}_{s3} \end{bmatrix}
\end{aligned}$$

其中

$$\begin{aligned}
f(k) &= \gamma^{-2k} \bar{z}^T(k+1) \bar{M}^T P_j \bar{M} \bar{z}(k+1) + \\
& \gamma^{-2k} \bar{e}^T(k) Q \bar{e}(k) + \\
& \gamma^{-2k} (\hat{u}(k) - U\hat{w}(k))^T R(\hat{u}(k) - U\hat{w}(k)) \\
\bar{\Phi}_{11} &= \gamma^{-2k-2l} \text{vecv}(z(k+l)) \\
\bar{\Phi}_{12} &= \text{vecv}(u(k+l)) \\
\bar{\Phi}_{13} &= 2\gamma^{-k-l} u^T(k+l) \otimes z^T(k+l)
\end{aligned}$$

可将式 (59) 转化为如下方程组

$$\bar{\psi}_j(k) [\text{vecs}(\bar{L}_{1j}); \text{vecs}(\bar{L}_{2j}); \text{vec}(\bar{L}_{3j})] = \bar{\varphi}_j(k) \quad (60)$$

当在线数据满足一定要求时, 上述方程组可由最小二乘方法求解. 如下引理给出了方程组 (60) 具有唯一解的条件.

引理 3. 方程组 (60) 可解并具有唯一解, 当且仅当

$$\begin{aligned}
\text{rank}(\bar{\psi}_j(k)) &= \frac{1}{2}((n_y + n_u)(n_x + n_w) + n_u) \times \\
& ((n_y + n_u)(n_x + n_w) + n_u + 1)
\end{aligned}$$

当引理 3 成立时, 方程组 (60) 可以由下式求

解, 为

$$\begin{aligned}
& [\text{vecs}(\bar{L}_{1j}); \text{vecs}(\bar{L}_{2j}); \text{vec}(\bar{L}_{3j})] = \\
& (\bar{\psi}_j^T(k) \bar{\psi}_j(k))^{-1} \bar{\psi}_j^T(k) \bar{\varphi}_j(k) \quad (61)
\end{aligned}$$

定义

$$\bar{P}_{j+1} = \bar{M}^T P_{j+1} \bar{M} \quad (62)$$

则 Riccati 方程 (57) 可由如下迭代公式求解

$$\bar{P}_{j+1} = \bar{L}_{1j} - \bar{L}_{3j} \bar{L}_{2j}^{-1} \bar{L}_{3j}^T \quad (63)$$

计算得到 \bar{P}_{j+1} 后, 将其代入 $\bar{\varphi}_j(k)$ 更新得到 $\bar{\varphi}_{j+1}(k)$, 继而可以更新方程 (60), 重复以上步骤可以得到序列 $\{\bar{P}_j\}_{j=0}^\infty$ 直至收敛, 对应的序列 $\{\bar{K}_j\}_{j=0}^\infty$ 为

$$\bar{K}_j = K_j \bar{M} - U \bar{M}_2 = \bar{L}_{2j}^{-1} \bar{L}_{3j}^T \quad (64)$$

至此, 基于输出反馈与强化学习的自适应最优输出调节算法如下.

算法 3. 基于输出反馈与强化学习的自适应最优输出调节算法

初始化: 选择任意的初始控制律 \bar{K}_0 , 终止条件常数 $\varepsilon > 0$, 半正定矩阵 \bar{P}_0 , $j \leftarrow 0$;

最优输出调节律在线计算: 利用如下迭代算法计算最优反馈增益, 在区间 $[k, k+s]$ 利用控制输入为 $\hat{u}(k) = -\bar{K}_0 \bar{z}(k) + n(k)$, 其中 $n(k)$ 为控制输入中添加的探测噪声, s 为使得引理 3 满足的数;

1) 利用式 (60) 计算得到 \bar{L}_{1j} , \bar{L}_{2j} , \bar{L}_{3j} ;

2) 利用式 (44) 计算 \bar{P}_{j+1} ;

3) 判断 $\|\bar{P}_{j+1} - \bar{P}_j\| < \varepsilon$ 是否成立, 如果成立则停止迭代, 并利用式 (64) 计算得到 \bar{K}_j , 反之则重复上述两步, 并令: $j \leftarrow j + 1$.

注 7. 算法 3 与算法 2 具有类似的特性, 其中 $(\bar{\psi}_j^T(k) \bar{\psi}_j(k))^{-1} \bar{\psi}_j^T(k)$ 在迭代过程中仅需要计算一次. \bar{K}_j 仅需要在 $\bar{M}^T P_{j+1} \bar{M}$ 收敛后计算一次. 该方法同样是一类 off-policy 方法, 可以保证计算结果是无偏的.

注 8. 本小节假设 U 是已知的, 该假设只需要在学习最优输出调节律时成立. 当 $B^T B$ 或 $S^T S$ 为非奇异矩阵时, $(\bar{B}^T P_j \bar{B} + S^T Q S)$ 是可逆的, 该情况下如果注 1 满足, 可将 R 设置为 0, 则 $\bar{\varphi}_j(k)$ 中的 $(\hat{u}(i) - U\hat{w}(i))^T R(\hat{u}(i) - U\hat{w}(i))$, $i = k, k+1, \dots, k+s$ 变为 0, 避免了 U 已知的要求.

4 算法收敛性与闭环稳定性分析

本节进行所设计的状态反馈与输出反馈自适应最优输出调节算法的收敛性分析与基于所设计的最优输出调节器的闭环系统稳定性分析, 如下两个定理分别给出了收敛性结论与稳定性结论.

定理 1. 当假设 1~3 成立, 引理 2 中条件满足时, 由算法 2 所得到的序列 $\{P_j\}_{j=0}^{\infty}$ 与 $\{K_j\}_{j=0}^{\infty}$ 最终会收敛至其最优值, 即 $\lim_{j \rightarrow \infty} P_j = P^*$, $\lim_{j \rightarrow \infty} K_j = K^*$. 另外, 当假设 1~4 成立, 引理 3 中条件满足时, 由算法 3 所得到的序列 $\{\bar{P}_j\}_{j=0}^{\infty}$ 与 $\{\bar{K}_j\}_{j=0}^{\infty}$ 最终会收敛至其最优值, 即 $\lim_{j \rightarrow \infty} \bar{P}_j = \bar{M}^T P^* \bar{M}$, $\lim_{j \rightarrow \infty} \bar{K}_j = K^* \bar{M} - U \bar{M}_2$.

证明. 当引理 2 条件满足时, 式 (43) 具有唯一解. 因此, 式 (43) 等价于算法 1 中的式 (30) 和式 (45) 等价于算法 1 中的式 (31), 这表明算法 2 的收敛性等价于算法 1 的收敛性. 基于引理 1, $\lim_{j \rightarrow \infty} P_j = P^*$, $\lim_{j \rightarrow \infty} K_j = K^*$ 得证. 对于算法 3, 当引理 3 条件满足时, 式 (63) 等价于算法 1 中的式 (30) 和式 (64) 等价于算法 1 中的式 (31), 基于引理 1, 可得 $\lim_{j \rightarrow \infty} \bar{P}_j = \bar{M}^T P^* \bar{M}$, $\lim_{j \rightarrow \infty} \bar{K}_j = K^* \bar{M} - U \bar{M}_2$.

定理 2. 考虑受扰动的线性离散系统 (1)~(2), 外部系统 (3)~(4), 当假设 1~4 成立时, 由算法 2 和算法 3 所得到的 K_j 与 \bar{K}_j 将使得闭环系统是渐近稳定的, 且跟踪误差 $e(k)$ 的收敛速率快于 γ^{-k} .

证明. 基于定理 1 的结论, 由算法 2 与算法 3 所得到的 K_j 与 \bar{K}_j 所控制的闭环对象为

$$\bar{x}(k+1) = (\bar{A} - \bar{B}K^*)\bar{x}(k) \quad (65)$$

$$\bar{e}(k) = C\bar{x}(k) \quad (66)$$

由于 K^* 是利用 Riccati 方程 (28) 求解得到, 因此 $\bar{A} - \bar{B}K^*$ 是 Schur 的, 所以有 $\lim_{k \rightarrow \infty} \bar{e}(k) = \lim_{k \rightarrow \infty} \gamma^k e(k) = 0$, 由此可得所计算的反馈控制增益 K_j 与 \bar{K}_j 解决了问题 1, 使得跟踪误差 $e(k)$ 的收敛速率快于 γ^{-k} .

5 仿真实验

本节进行所提算法的仿真实验研究, 首先介绍仿真实验对象与实验参数, 之后分别进行基于状态反馈的仿真实验与基于输出反馈的仿真实验.

5.1 仿真实验对象与实验参数

考虑如下受扰动的线性离散时间系统

$$x(k+1) = \begin{bmatrix} 0 & 1 \\ -1 & -3 \end{bmatrix} x(k) + \begin{bmatrix} 0 \\ 0.6 \end{bmatrix} u(k) + w(k)$$

$$y(k) = [1 \ 0] x(k) + u(k)$$

对应的外部系统与参考信号为

$$w(k+1) = \begin{bmatrix} \cos(0.2) & \sin(0.2) \\ -\sin(0.2) & \cos(0.2) \end{bmatrix} w(k)$$

$$y_d(k) = [1 \ 0] w(k)$$

问题 1 中的矩阵参数选择为 $Q = R = 1$, 问题 2 中的矩阵参数选择为 $M = I$, 收敛速率 $\gamma = 1.2$, 利用式 (9) 求解输出调节方程 (7)~(8) 的解 X 和 U 为

$$X = \begin{bmatrix} 0.8506 & 0.066 \\ -0.1795 & 0.2337 \end{bmatrix}, U = [0.1494 \ -0.066]$$

则最优的 P^* , K^* , L^* 与 \bar{K}^* 分别为

$$P^* = \begin{bmatrix} 8.8818 & 16.1083 \\ 16.1083 & 32.1106 \end{bmatrix}$$

$$K^* = [-1.4343 \ -3.7173]$$

$$L^* = [-0.4032 \ -1.0293]$$

$$\bar{K}^* = \begin{bmatrix} -15.8383 & 31.2417 & -6.3175 & -10.985 \\ 13.1619 & -22.8457 & -6.3697 & 17.5763 \end{bmatrix}$$

5.2 基于状态反馈的仿真实验

本小节进行基于状态反馈的仿真实验, 仿真实验中, 初始控制律 $K_0 = [-1 \ -3]^T$, 终止条件常数 $\varepsilon = 0.001$, 半正定矩阵 $P_0 = 0$, 矩阵序列

$$\begin{bmatrix} X_0 \\ U_0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} X_1 \\ U_1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix},$$

$$\begin{bmatrix} X_2 \\ U_2 \end{bmatrix} = \begin{bmatrix} -\sqrt{2} & 0 \\ 0 & 0 \\ \sqrt{2} & 0 \end{bmatrix}, \begin{bmatrix} X_3 \\ U_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\sqrt{2} & 2 \\ 0 & -1 \end{bmatrix},$$

$$\begin{bmatrix} X_4 \\ U_4 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} X_5 \\ U_5 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ -\sqrt{2} & 2 \\ 0 & 1 \end{bmatrix}.$$

探测噪声 $n(k)$ 为白噪声, 被控对象的初始状态为 $x(1) = [1 \ 2]^T$ 与 $w(1) = [2 \ 1]^T$. 由引理 2 可知, 求解公式 (42) 至少需要 15 组数据, 故 s 需大于 14, 仿真实验中选择 $s = 17$.

仿真结果如图 1~3 所示, 图 1 表示基于状态反馈的输出 $y(k)$ 与参考信号 $y_d(k)$ 的轨迹, 由该图可知本文所提方法能够在系统矩阵 A, B, D, E 未知时实现自适应输出调节, 图 2 表示基于状态反馈的 $\|P_j - P^*\|$ 与 $\|K_j - K^*\|$ 的误差轨迹, 由图可知经过 13 步迭代算法收敛, 图 3 表示基于状态反馈的误差 $e(k)$ 与 $\gamma^{-k}e(k_0)$ 的对比曲线, 实验结果表明所设计的控制器能够使得跟踪误差收敛快于 γ^{-k} .

5.3 基于输出反馈的仿真实验

本小节进行基于状态反馈的仿真实验, 仿真实验中, 初始控制律

$$\bar{K}_0 = \begin{bmatrix} -13.5899 & 24.4082 & -4.4063 & -8.4499 \\ 11.4299 & -17.4962 & -5.6248 & 13.5199 \end{bmatrix}$$

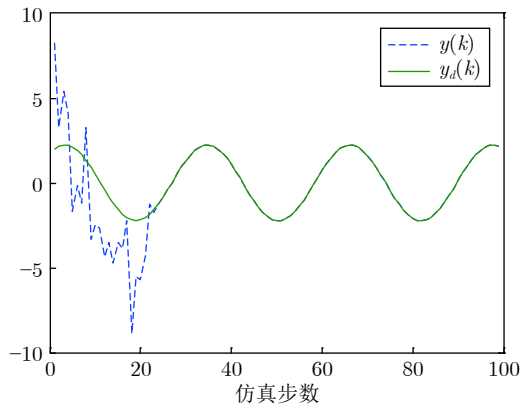


图 1 基于状态反馈的输出 $y(k)$ 与参考信号 $y_d(k)$ 轨迹
Fig.1 Trajectories of the output $y(k)$ and the reference signal $y_d(k)$ via state feedback

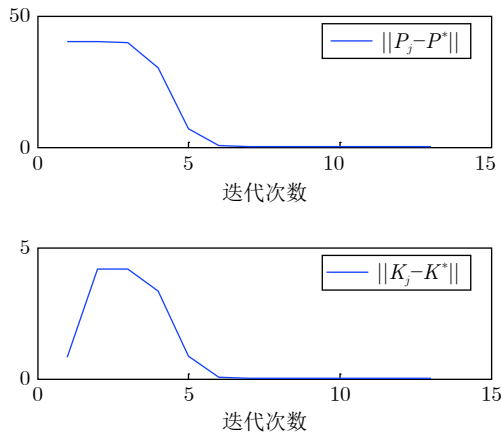


图 2 基于状态反馈的 $\|P_j - P^*\|$ 与 $\|K_j - K^*\|$ 误差轨迹
Fig.2 Trajectory of the error between $\|P_j - P^*\|$ and $\|K_j - K^*\|$ via state feedback

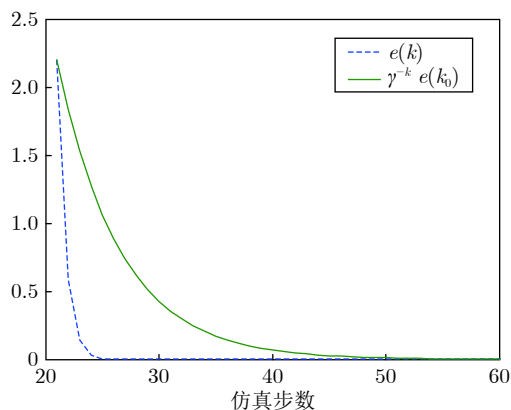


图 3 基于状态反馈的误差 $e(k)$ 与 $\gamma^{-k}e(k_0)$ 对比曲线
Fig.3 Comparison curve of $e(k)$ and $\gamma^{-k}e(k_0)$ via state feedback

终止条件常数 $\varepsilon = 80$, 半正定矩阵 $P_0 = 0$, 探测噪声 $n(k)$ 为白噪声, 被控对象的初始状态为 $x(1) = [1 \ 2]^T$ 与 $w(1) = [2 \ 1]^T$. 由引理 3 可知, 求解式 (60) 至少需要 45 组数据, 故 s 需大于 44, 仿真实验中选择 $s = 64$.

仿真结果如图 4 ~ 6 所示, 图 4 表示基于 s 输出反馈的输出 $y(k)$ 与参考信号 $y_d(k)$ 的轨迹, 由该图可知本文所提方法能够实现自适应输出调节, 图 5 表示基于输出反馈的 $\|\bar{P}_j - \bar{P}^*\|$ 与 $\|\bar{K}_j - \bar{K}^*\|$ 的误差轨迹, 图 6 表示基于输出反馈的误差 $e(k)$ 与 $\gamma^{-k}e(k_0)$ 的对比曲线, 实验结果表明所设计的控制器能够使得跟踪误差收敛快于 γ^{-k} .

5.4 对比仿真实验

本小节进行对比仿真实验, 其中对比方法选用

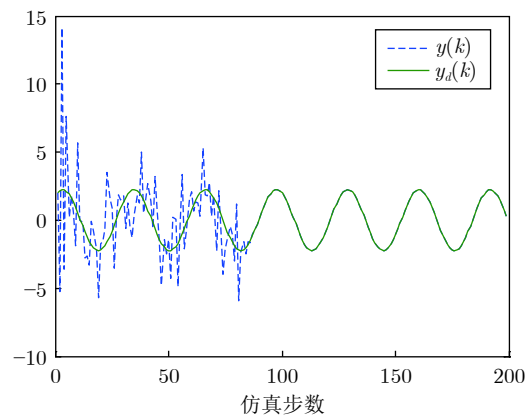


图 4 基于输出反馈的输出 $y(k)$ 与参考信号 $y_d(k)$ 轨迹
Fig.4 Trajectories of the output $y(k)$ and the reference signal $y_d(k)$ via output feedback

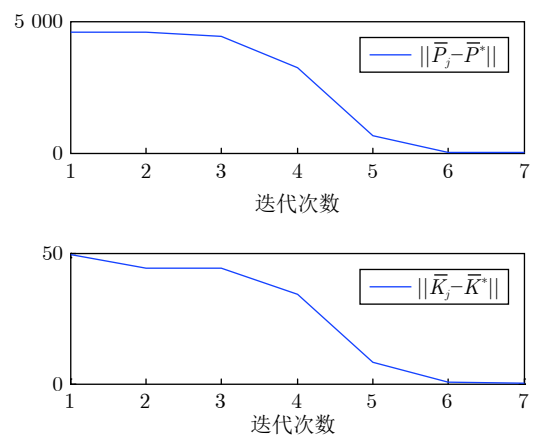


图 5 基于输出反馈的 $\|\bar{P}_j - \bar{P}^*\|$ 与 $\|\bar{K}_j - \bar{K}^*\|$ 误差轨迹

Fig.5 Trajectory of the error between $\|\bar{P}_j - \bar{P}^*\|$ and $\|\bar{K}_j - \bar{K}^*\|$ via output feedback

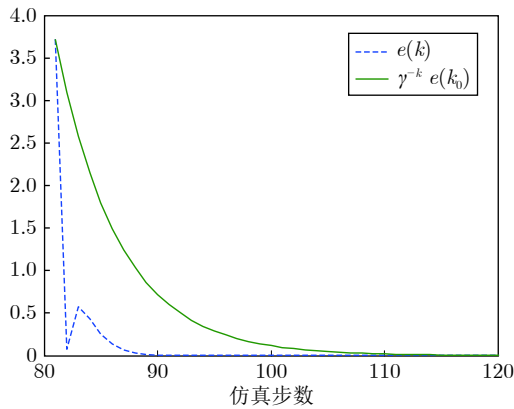
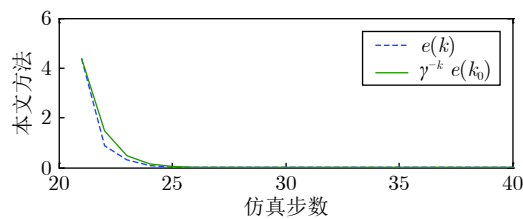


图 6 基于输出反馈的误差 $e(k)$ 与 $\gamma^{-k}e(k_0)$ 对比曲线

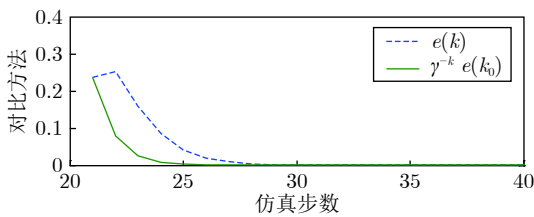
Fig.6 Comparison curve of $e(k)$ and $\gamma^{-k}e(k_0)$ via output feedback

文献 [24] 方法, 对比实验的参数选择为 $Q = 1$, $R = 30$, 收敛速率 $\gamma = 3$. 由于文献 [24] 中的方法无法求解输出调节方程 (7) ~ (8) 的解 X 和 U , 对比实验中求解 X 和 U 均使用本文的方法. 对比方法中的初始控制策略为稳定的. 对比仿真结果如图 7 表示, 实验结果表明, 与对比方法相比, 在相同的权重矩阵参数下, 本文所设计的控制器使得跟踪误差收敛快于 γ^{-k} , 而对比方法计算得到的控制器使得跟踪误差收敛慢于 γ^{-k} .



(a) 基于本文方法的误差曲线 $e(k)$

(a) Error curve $e(k)$ via the developed approach in this paper



(b) 基于文献[24]方法的误差曲线 $e(k)$

(b) Error curve $e(k)$ via the approach in [24]

图 7 对比仿真结果

Fig.7 Comparison of simulation results

6 结束语

本文针对具有未知动态与收敛速率要求的受扰离散线性系统的输出调节问题, 提出了基于状态反馈与输出反馈的自适应最优输出调节算法, 该算法

不需要稳定的初始控制律与部分模型知识, 利用在线算法求解得到最优的输出调节器, 同时还能够保证跟踪误差的收敛速率满足预先给定的要求. 本文的后续工作将着重于研究基于动态反馈的输出调节算法, 以克服对部分模型知识的要求.

References

- 1 Åström K J, Tore H. *PID Controllers: Theory, Design, and Tuning*. Research Triangle Park, NC: Instrument Society of America, 1995.
- 2 Garcia C E, Prett D M, Morari M. Model predictive control: theory and practice: A survey. *Automatica*, 1989, **25**(3): 335-348
- 3 Francis B A. The linear multivariable regulator problem. *SIAM Journal on Control and Optimization*, 1977, **15**(3): 486-505
- 4 Isidori A, Byrnes C I. Output regulation of nonlinear systems. *IEEE Transactions on Automatic Control*, 1990, **35**(2): 131-140
- 5 Ding Z T. Output regulation of uncertain nonlinear systems with nonlinear exosystems. *IEEE Transactions on Automatic Control*, 2006, **51**(3): 498-503
- 6 Huang J, Chen Z. A general framework for tackling the output regulation problem. *IEEE Transactions on Automatic Control*, 2004, **49**(12): 2203-2218
- 7 Parks P. Liapunov redesign of model reference adaptive control systems. *IEEE Transactions on Automatic Control*, 1966, **11**(3): 362-367
- 8 Tian Tao-Tao, Hou Zhong-Sheng, Liu Shi-Da, Deng Zhi-Dong. Model-free adaptive control based lateral control of self-driving car. *Acta Automatica Sinica*, 2017, **43**(11): 1931-1940 (田涛涛, 侯忠生, 刘世达, 邓志东. 基于无模型自适应控制的无人驾驶汽车横向控制方法. *自动化学报*, 2017, **43**(11): 1931-1940)
- 9 Yu Xin-Bo, He Wei, Xue Cheng-Qian, Sun Yong-Kun, Sun Chang-Yin. Disturbance observer-based adaptive neural network tracking control for robots. *Acta Automatica Sinica*, 2019, **45**(7): 1307-1324 (于欣波, 贺威, 薛程谦, 孙永坤, 孙长银. 基于扰动观测器的机器人自适应神经网络跟踪控制研究. *自动化学报*, 2019, **45**(7): 1307-1324)
- 10 Modares H, Lewis F L. Linear quadratic tracking control of partially-unknown continuous-time systems using reinforcement learning. *IEEE Transactions on Automatic Control*, 2014, **59**(11): 3051-3056
- 11 Xue W Q, Fan J L, Lopez V G, Jiang Y, Chai T Y, Lewis F L. Off-policy reinforcement learning for tracking in continuous-time systems on two time-scales. *IEEE Transactions on Neural Networks and Learning Systems*, 2021, **32** (10), 4334-4346
- 12 Kiumarsi B, Lewis F L, Modares H, Karimpour A, Naghibisistani M B. Reinforcement Q-learning for optimal tracking control of linear discrete-time systems with unknown dynamics. *Automatica*, 2014, **50**(4): 1167-1175
- 13 Jiang Y, Fan J, Chai T, Lewis F L, Li J N. Tracking control for linear discrete-time networked control systems with unknown dynamics and dropout. *IEEE Transactions on Neural Networks and Learning System*, 2018, **29**(10): 4607-4620
- 14 Wu Qian, Fan Jia-Lu, Jiang Yi, Chai Tian-You. Data-driven dual-rate control for mixed separation thickening process in a wireless network environment. *Acta Automatica Sinica*, 2019, **45**(6): 1122-1135 (吴倩, 范家璐, 姜艺, 柴天佑. 无线网络环境下数据驱动混合选别浓密过程双率控制方法. *自动化学报*, 2019, **45**(6): 1122-1135)
- 15 Xue W Q, Fan J L, Lopez V G, Li J N, Jiang Y, Chai T Y, Lewis F L. New methods for optimal operational control of industrial processes using reinforcement learning on two time scales. *IEEE Transactions on Industrial Informatics*, 2020, **16**(5): 3085-3099
- 16 Modares H, Lewis F L. Optimal tracking control of nonlinear partially-unknown constrained-input systems using integral reinforcement learning. *Automatica*, 2014, **50**(7): 1780-1792
- 17 Kiumarsi B, Lewis F L. Actor-critic-based optimal tracking for partially unknown nonlinear discrete-time systems. *IEEE Transactions on Neural Networks and Learning Systems*, 2014, **26**(1): 140-151
- 18 Jiang Y, Fan J L, Chai T Y, Li J N, Lewis F L. Data-driven

flotation industrial process operational optimal control based on reinforcement learning. *IEEE Transactions on Industrial Informatics*, 2018, **14**(5): 1974–1989

19 Jiang Y, Fan J L, Chai T Y, Lewis F L. Dual-rate operational optimal control for flotation industrial process with unknown operational model. *IEEE Transactions on Industrial Electronics*, 2019, **66**(6): 4587–4599

20 Gao W N, Jiang Z P. Adaptive dynamic programming and adaptive optimal output regulation of linear systems. *IEEE Transactions on Automatic Control*, 2016, **61**(12): 4164–4169

21 Gao W N, Jiang Z P, Lewis F L, Wang Y B. Leader-to-formation stability of multi-agent systems: an adaptive optimal control approach. *IEEE Transactions on Automatic Control*, 2018, **63**(10): 3581–3587

22 Chen C, Modares H, Xie K, Lewis F L, Wan Y, Xie S L. Reinforcement learning-based adaptive optimal exponential tracking control of linear systems with unknown dynamics. *IEEE Transactions on Automatic Control*, 2019, **64**(11): 4423–4438

23 Chen C, Lewis F L, Xie K, Xie S L, Liu Y L. Off-policy learning for adaptive optimal output synchronization of heterogeneous multi-agent systems. *Automatica*, 2020, **119**: 109081

24 Jiang Y, Kiumarsi B, Fan J L, Chai T Y, Li J N, Lewis F L. Optimal output regulation of linear discrete-time systems with unknown dynamics using reinforcement learning. *IEEE Transactions on Cybernetics*, 2020, **50**(7): 3147–3156

25 Pang Wen-Yan, Fan Jia-Lu, Jiang Yi, Lewis Frank Leroy. Optimal output regulation of partially linear discrete-time systems using reinforcement learning. *Acta Automatica Sinica*, DOI: 10.16383/j.aas.c190853 (庞文砚, 范家璐, 姜艺, 刘易斯·弗兰克. 基于强化学习的部分线性离散时间系统最优输出调节. *自动化学报*, DOI: 10.16383/j.aas.c190853)

26 Fan J L, Wu Q, Jiang Y, Chai T Y, Lewis F L. Model-free optimal output regulation for linear discrete-time lossy networked control systems. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2020, **50**(11): 4033–4042

27 Gao W N, Jiang Z P. Learning-based adaptive optimal tracking control of strict-feedback nonlinear systems. *IEEE Transactions on Neural Networks and Learning System*, 2018, **29**(6): 2614–2624

28 Jiang Y, Fan J L, Gao W N, Chai T Y, Lewis F L. Cooperative adaptive optimal output regulation of discrete-time nonlinear multi-agent systems. *Automatica*, 2020, **121**: 109149

29 Kiumarsi B, Lewis F L, Modares H, Karimpour A, Naghibisistani M B. Optimal tracking control of unknown discrete-time linear systems using input-output measured data. *IEEE Transactions on Cybernetics*, 2015, **45**(12): 2770–2779

30 Gao W N, Jiang Z P. Adaptive optimal output regulation of time-delay systems via measurement feedback. *IEEE Transactions on Neural Networks and Learning System*, 2018, **30**(3): 938–945

31 Zhang Chun-Yan, Qi Guo-Qing, Li Yin-Ya, Sheng An-Dong. Standoff tracking control with respect to moving target via finite-time stabilization. *Acta Automatica Sinica*, 2018, **44**(11): 2056–2067 (张春燕, 戚国庆, 李银伢, 盛安冬. 一种基于有限时间稳定的环绕控制器设计. *自动化学报*, 2018, **44**(11): 2056–2067)

32 Hong Y G, Xu Y S, Huang J. Finite-time control for robot manipulators. *Systems and control letters*, 2002, **46**(4): 243–253

33 Huang J. *Nonlinear Output Regulation: Theory and Applications*. SIAM, 2004.

34 Krener A J. *The construction of optimal linear and nonlinear regulators*. Systems, Models and Feedback: Theory and Applications. Springer, 1992.

35 Arnold W F, Laub A J. Generalized eigen problem algorithms and software for algebraic Riccati equations. *Proceedings of the IEEE*, 1984, **72**(12): 1746–1754

36 Lewis F L, Vrabie D, Syrmos V L. *Optimal Control*. John Wiley & Sons, 2012.

37 Lancaster P, Rodman L. *Algebraic Riccati Equations*. New York: Oxford University Press, 1995.

38 Hewer G. An iterative technique for the computation of the steady state gains for the discrete optimal regulator. *IEEE Transactions on Automatic Control*, 1971, **16**(4): 382–384

39 Li J N, Chai T Y, Lewis F L, Ding Z T, Jiang Y. Off-policy Interleaved Q-learning: Optimal control for affine nonlinear discrete-time systems. *IEEE Transactions on Neural Networks and Learning System*, 2019, **30**(5): 1308–1320

40 Kiumarsi B, Lewis F L, Jiang Z P. H_∞ control of linear discrete-time systems: Off-policy reinforcement learning. *Automatica*, 2017, **78**: 144–152

41 Li Zhen, Fan Jia-Lu, Jiang Yi, Chai Tian-You. A model-free H_∞ method based on off-policy with output data feedback. *Acta Automatica Sinica*, 2021, **47**(9), 2182–21932 (李臻, 范家璐, 姜艺, 柴天佑. 一种基于Off-policy的无模型输出数据反馈 H_∞ 控制方法. *自动化学报*, 2021, **47**(9), 2182–2193)



姜艺 2020年获东北大学控制理论与控制工程博士学位。主要研究方向为工业过程运行控制,网络控制,自适应控制,网络控制,自适应动态规划和强化学习。

E-mail: yjian22@cityu.edu.hk

(JIANG Yi) He received his Ph. D.

degree in Control Theory and Engineering from Northeastern University in 2020. His research interest covers industrial process operational control, networked control, adaptive dynamic programming and reinforcement learning.)



范家璐 东北大学流程工业综合自动化国家重点实验室副教授。2011年获浙江大学控制科学与工程系博士学位,2009至2010年美国宾夕法尼亚州立大学访问学者。主要研究方向为工业过程运行控制,工业无线传感器网络与移动社会网络。

E-mail: jlfan@mail.neu.edu.cn

(FAN Jia-Lu) Associate professor at the State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University. She received her Ph. D. degree from Zhejiang University in 2011. She was a visiting scholar at Pennsylvania State University during 2009 to 2010. Her research interest covers networked operational control, industrial wireless sensor networks and mobile social networks.)



柴天佑 中国工程院院士,东北大学教授。1985年获东北大学博士学位。主要研究方向为自适应控制,智能解耦控制,流程工业综合自动化理论和方法与技术。本文通信作者。

E-mail: tychai@mail.neu.edu.cn

(CHAI Tian-You) Academician of Chinese Academy of Engineering, professor at Northeastern University. He received his Ph. D. degree from Northeastern University in 1985. His research interest covers adaptive control, intelligent decoupling control and integrated automation theory, method and technology of industrial process. Corresponding author of this paper.)