

控制与决策

Control and Decision

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

高伟男, 杨涛, 柴天佑

引用本文:

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.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2023.0335>

您可能感兴趣的其他文章

Articles you may be interested in

[基于数据驱动的非线性网络系统自适应迭代学习控制](#)

Data driven adaptive learning control of nonlinear network system

控制与决策. 2021, 36(6): 1523–1528 <https://doi.org/10.13195/j.kzyjc.2019.1182>

[一种改进的紧格式无模型自适应控制方法](#)

An improved compact form model free adaptive control method

控制与决策. 2021, 36(2): 436–442 <https://doi.org/10.13195/j.kzyjc.2019.0635>

[线控转向系统的自适应高阶滑模控制](#)

Adaptive higher-order sliding mode control for SbW system

控制与决策. 2021, 36(6): 1529–1536 <https://doi.org/10.13195/j.kzyjc.2019.1526>

[磁悬浮开关磁阻电机的自适应终端滑模控制](#)

Adaptive terminal sliding mode control of bearingless switched reluctance motor

控制与决策. 2021, 36(6): 1449–1456 <https://doi.org/10.13195/j.kzyjc.2019.1064>

[航天器输入受限的鲁棒自适应姿态跟踪控制](#)

Robust adaptive attitude tracking control of spacecraft with constrained inputs

控制与决策. 2021, 36(9): 2297–2304 <https://doi.org/10.13195/j.kzyjc.2020.0013>

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

高伟男[†], 杨涛, 柴天佑

(东北大学 流程工业综合自动化国家重点实验室, 沈阳 110819)

摘要: 当设计前馈-反馈控制器来实现连续时间线性系统的自适应最优输出调节时, 往往需要依赖调节器方程的解, 然而在系统动态不确定的前提下, 获得这个方程的解一般非常困难. 现有的自适应最优输出调节方法从选取调节方程解的基矩阵出发, 通过在线数据学习每个基矩阵所对应的权重. 但是, 当被控对象或外部系统阶数过高时, 这种方法的计算复杂度亦非常高. 对此, 提出一种新的无模型方法来实现闭环系统的自适应最优输出调节, 调节器方程的解由梯度下降法逼近. 需要指出的是, 该方法通过在线数据而不是模型信息来获得梯度. 仿真结果验证了所提出方法的有效性.

关键词: 输出调节; 自适应动态规划; 梯度下降法; 自适应控制; 优化控制

中图分类号: TP13 **文献标志码:** A

DOI: 10.13195/j.kzyjc.2023.0335

引用格式: 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.

Adaptive optimal output regulation based on adaptive dynamic programming and gradient descent method

GAO Wei-nan[†], YANG Tao, CHAI Tian-you

(State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang 110819, China)

Abstract: When dealing with adaptive optimal output regulation problems of linear continuous-time systems, the development of an optimal feedback-feedforward controller usually relies on the solution to regulator equations. However, it is nontrivial to obtain this solution when the system dynamic is uncertain. Existing studies of adaptive optimal output regulation often begin with selecting basis of solution to regulator equations and then use online data to learn the weights. But this method is computationally expensive if the order of the plant or the exosystem is high. In this paper, we propose a novel model-free solution to adaptive optimal output regulation problems where the solution to regulator equations can be approximated via gradient descent methods. Notably, the gradient can be learned by online data instead of using model knowledge. Simulation results demonstrate the effectiveness of the proposed method.

Keywords: output regulation; adaptive dynamic programming; gradient descent method; adaptive control; optimal control

0 引言

作为控制理论中的一个经典问题, 输出调节问题^[1-3]旨在设计反馈控制器以实现闭环系统扰动抑制与输出渐近跟踪. 然而在实际工程应用中, 控制器设计的目标不仅要保证闭环系统实现输出调节, 还需在系统动态未知的情况下尽可能地优化其暂态性能, 这被称为自适应最优输出调节问题.

自适应动态规划^[4-9]是一种针对不确定动态系

统的自适应最优控制方法. 该方法的优越性在于其可以利用在线数据通过学习来逼近动态系统的最优控制器. 换言之, 自适应动态规划可以通过数据信息来补偿设计者对系统动态的未知性. 基于该优势, 文献[10-14]将自适应动态规划与输出调节理论结合起来探索自适应最优输出调节问题的解决方法. 虽然围绕自适应最优输出调节问题已经取得了很好进展, 但是, 依然存在两个关键问题亟待解决. 1) 欲实现输

收稿日期: 2023-03-23; 录用日期: 2023-05-10.

基金项目: 国家自然科学基金重大项目(61991404, 61991403); 2020年度辽宁省科技重大专项计划项目(2020JH1/10100008).

责任编辑: 叶丹.

[†]通讯作者. E-mail: gaown@mail.neu.edu.cn.

出调节往往需求解调节器方程. 文献[10]尝试找到调节器方程解的基矩阵, 通过在线数据学习每个基矩阵的西尔韦斯特映射, 再通过学习到的映射来确定调节器方程解. 虽然这种方法可以在线求解调节器方程, 但当系统维数过高时, 所需的基矩阵个数过多, 从而导致该方法的计算复杂度过高. 因此, 在系统动态未知的情况下, 如何有效地求解调节器方程仍是一个技术上的难题. 2) 现有自适应最优输出调节问题的研究^[10-14]往往需要已知部分系统矩阵的信息. 如何在系统所有矩阵信息均未知的情况下设计自适应最优控制器仍需深入探索.

针对以上两个关键问题, 本文将自适应动态规划与梯度下降法相结合, 提出一种新的自适应最优前馈-反馈控制器设计方法来实现输出调节. 首先基于调节器方程建立一个二次型指标函数, 通过梯度下降方法来寻找指标函数的最小值, 从而求解调节器方程. 其中一个技术难点是, 如何在系统模型未知的情况下得到指标函数的梯度信息. 为克服此难点, 本文基于在线数据学习得到梯度信息, 并从理论上证明所学梯度信息与实际梯度等价. 在系统所有矩阵未知的情况下, 本文提出的方法可以同时逼近黎卡提方程和调节器方程的解析解, 从而学习系统的自适应最优前馈和反馈增益.

在本文中: \mathbf{Z}_+ 表示非负整数集合; $\|\cdot\|$ 为计算向量欧氏范数的运算符; 对于任意矩阵 $A \in \mathbf{R}^{m \times n}$, $\ker A$ 是矩阵 A 的核; 当 A 为方阵时, $\sigma(A)$ 是方阵 A 所有特征值组成的集合; \otimes 是克罗内克积算子; $\text{vec}(B) = [b_1^T, b_2^T, \dots, b_m^T]^T$ 是将矩阵 $B \in \mathbf{R}^{r \times m}$ 向量化, 其中 $b_i \in \mathbf{R}^r$ 是矩阵 B 的第 i 列. 相应地, 针对一个向量 $c \in \mathbf{R}^{nq}$, 可通过 $C = \text{vec}_{n \times q}^{-1}(c)$ 得到一个矩阵 $C \in \mathbf{R}^{m \times q}$ 满足 $c = \text{vec}(C)$. 针对任意对称方阵 $D \in \mathbf{R}^{m \times m}$, $\text{vecs}(D) = [d_{11}, 2d_{12}, \dots, 2d_{1m}, d_{22}, 2d_{23}, \dots, 2d_{m-1,m}, d_{mm}]^T \in \mathbf{R}^{\frac{1}{2}m(m+1)}$. 针对任意列向量 $v \in \mathbf{R}^n$, $\text{vecv}(v) = [v_1^2, v_1v_2, \dots, v_1v_n, v_2^2, v_2v_3, \dots, v_{n-1}v_n, v_n^2]^T \in \mathbf{R}^{\frac{1}{2}n(n+1)}$.

1 最优输出调节问题描述与求解

本文考虑如下—类线性系统:

$$\dot{v} = Sv, \quad (1)$$

$$\dot{x} = Ax + Bu + Ev, \quad (2)$$

$$e = Cx + Fv. \quad (3)$$

其中: $x \in \mathbf{R}^n$ 是状态向量, $u \in \mathbf{R}^m$ 是输入向量, $e \in \mathbf{R}^m$ 是跟踪误差. 自治系统(1)称为外部系统, 该系统可同时生成扰动信号和参考信号. 自治系统的状态

向量 $v \in \mathbf{R}^q$ 也被称为外部状态. 系统(1)~(3)中矩阵的维数分别为 $A \in \mathbf{R}^{n \times n}$, $B \in \mathbf{R}^{n \times m}$, $C \in \mathbf{R}^{m \times n}$, $E \in \mathbf{R}^{n \times q}$, $F \in \mathbf{R}^{m \times q}$, $S \in \mathbf{R}^{q \times q}$.

针对线性系统(1)~(3), 本文给出以下几个基本假设.

假设1 (A, B) 是可镇定的.

假设2 矩阵 S 的所有特征根都不重复地分布在虚轴上.

假设3 以下满秩条件成立:

$$\text{rank} \begin{bmatrix} A - \lambda I & B \\ C & 0 \end{bmatrix} = n + m, \quad \forall \lambda \in \sigma(S).$$

注1 假设1~假设3均为现有文献^[1,10]解决输出调节问题的基本假设. 假设2保证外部系统(1)是临界稳定的, 外部状态亦不会发散.

欲解决输出调节问题, 本质上需要设计控制器使得闭环系统满足两个条件: 1) 当外部状态 $v(t)$ 恒等于零时, 闭环系统可在原点处指数稳定; 2) 针对任意初值 $x(0) \in \mathbf{R}^n$, $v(0) \in \mathbf{R}^q$, 系统跟踪误差 $e(t)$ 渐近地收敛到零, 即实现渐近跟踪.

调节器方程是求解输出调节问题的一个重要媒介. 针对线性输出调节问题, 调节器方程是一个西尔韦斯特方程, 即

$$0 = XS - AX - BU - E := f_1(X, U), \quad (4)$$

$$0 = CX + F := f_2(X). \quad (5)$$

当假设3满足时, 调节器方程(4)和(5)存在唯一解 $(X^*, U^*) \in \mathbf{R}^{n \times q} \times \mathbf{R}^{m \times q}$. 如能选择一个反馈增益 $K \in \mathbf{R}^{m \times n}$ 使得矩阵 $A - BK$ 是一个赫尔维茨矩阵, 则可根据所选择的反馈增益和调节器方程的解设计一个前馈增益 $L = KX^* + U^*$, 从而获得如下前馈-反馈控制器:

$$u(t) = -Kx(t) + Lv(t). \quad (6)$$

在假设1~假设3满足的前提下, 文献[1]已证明, 应用前馈-反馈控制器(6)的闭环系统可以实现输出调节. 并且对于任意的时间 $t \in \mathbf{R}_+$, 如果将信号 $\tilde{x}(t)$ 和 $\tilde{u}(t)$ 分别定义为 $\tilde{x}(t) = x(t) - X^*v(t)$ 和 $\tilde{u}(t) = u(t) - U^*v(t)$, 则可获知 $\lim_{t \rightarrow \infty} \tilde{x}(t) = 0$, $\lim_{t \rightarrow \infty} \tilde{u}(t) = 0$.

相比于传统输出调节, 最优输出调节问题还需考虑优化系统的暂态响应. 更严格地讲, 解决最优输出调节问题, 需要设计一个最优的前馈-反馈控制器, 即

$$u(t) = -K^*x(t) + L^*v(t), \quad (7)$$

$$L^* = K^*X^* + U^*, \quad (8)$$

使得闭环系统实现输出调节, 并且其反馈控制增益 K^* 可通过求解以下动态规划问题获得:

$$\min_{\tilde{u}} \int_0^\infty (\tilde{x}^T(t)Q\tilde{x}(t) + \tilde{u}^T(t)R\tilde{u}(t))dt; \quad (9)$$

$$\text{s.t. } \dot{\tilde{x}}(t) = A\tilde{x}(t) + B\tilde{u}(t). \quad (10)$$

其中 Q 是正定对称矩阵. 系统(10)可认为是系统(2)的误差系统.

不难看出, 动态规划问题(9)和(10)是一个线性二次型调节器问题. 根据最优控制理论, 该问题下的最优控制器为 $\tilde{u}(t) = -K^*\tilde{x}(t)$, 其最优反馈增益为

$$K^* = R^{-1}B^T P^*. \quad (11)$$

对称正定矩阵 P^* 可通过求解如下代数黎卡提方程获得:

$$A^T P + PA + Q - PBR^{-1}B^T P = 0. \quad (12)$$

至此, 当所有系统动态均精确已知时, 即可通过求解调节器方程(4)、(5)和代数黎卡提方程(12)来设计最优的前馈-反馈控制器(7)以实现最优输出调节.

2 自适应最优输出调节问题描述与求解

针对线性系统(1)~(3), 自适应最优输出调节问题旨在当系统动态不确定时, 寻找最优前馈-反馈控制器(7). 本节将讨论如何通过结合策略迭代和梯度下降法学习这个最优控制器.

通过以上描述可知, 探索(自适应)最优输出调节的关键是求解代数黎卡提方程(12)以及调节器方程(4)和(5). 由于代数黎卡提方程(12)是一个非线性方程, 当系统维数较高时, 求解起来比较复杂. 为降低其复杂度, Kleinman^[15]提出了一种策略迭代方法, 即给定一个初始可接受的反馈控制增益 K_0 (使闭环系统稳定并且所需成本有限), 通过策略迭代方法可以不断更新反馈控制增益. 例如, 已知一个反馈增益 K_k 使得 $A - BK_k$ 是一个赫尔维茨矩阵, 策略迭代即是通过求解如下李雅普诺夫方程来更新系统的价值 P_k :

$$0 = (A - BK_k)^T P_k + P_k(A - BK_k) + Q + K_k^T R K_k, \quad (13)$$

并通过下式更新反馈增益 K_{k+1} :

$$K_{k+1} = R^{-1}B^T P_k. \quad (14)$$

文献[15]已证明, 通过策略迭代方法获得的每一代反馈控制增益都是可接受的, 并且保证了收敛性, 即 $\lim_{k \rightarrow \infty} P_k = P^*$, $\lim_{k \rightarrow \infty} K_k = K^*$.

调节器方程也可以通过迭代的方法进行逼近. 首先定义两个向量 $z^x = \text{vec}(X)$ 和 $z^u = \text{vec}(U)$. 其次, 基于调节器方程(4)和(5), 定义如下两个函数:

$$\mathcal{R}_1(z^x, z^u) = \text{vec}(f_1(X, U)) = \text{vec}(XS - AX - BU - E)$$

$$[(S^T \otimes I_n) - (I_q \otimes A)]z^x - (I_q \otimes B)z^u - \text{vec}(E), \quad (15)$$

$$\mathcal{R}_2(z^x) = \text{vec}(f_2(X)) = \text{vec}(CX + F) = (I_q \otimes C)z^x + \text{vec}(F). \quad (16)$$

为了求解调节器方程, 可建立如下无约束凸优化问题:

$$\min_{z^x, z^u} J(z^x, z^u) = \|\mathcal{R}_1(z^x, z^u)\|^2 + \|\mathcal{R}_2(z^x)\|^2. \quad (17)$$

基于梯度下降法, 若选取足够小的步长 $\epsilon_x, \epsilon_u > 0$, 则可通过迭代下式来逼近调节器方程的解:

$$\begin{aligned} z_{k+1}^x &= z_k^x - \epsilon_x \frac{\partial J}{\partial z^x} \Big|_{(z_k^x, z_k^u)} = \\ & z_k^x - 2\epsilon_x \frac{\partial \mathcal{R}_1^T}{\partial z^x} \mathcal{R}_1(z_k^x) - 2\epsilon_x \frac{\partial \mathcal{R}_2^T}{\partial z^x} \mathcal{R}_2(z_k^x), \\ z_{k+1}^u &= z_k^u - \epsilon_u \frac{\partial J}{\partial z^u} \Big|_{(z_k^x, z_k^u)} = \\ & z_k^u - 2\epsilon_u \frac{\partial \mathcal{R}_1^T}{\partial z^u} \mathcal{R}_1(z_k^x). \end{aligned} \quad (18)$$

然而, 无论是利用策略迭代求解黎卡提方程, 还是利用梯度下降法求解调节器方程, 都不免要依赖系统的动态信息, 这与实现自适应输出调节的初衷相违背. 因此, 本文尝试在系统动态未知的情况下, 通过在在线数据信息来求解这个两个方程.

针对任意一对矩阵 $(X_k, U_k) \in \mathbf{R}^{n \times q} \times \mathbf{R}^{m \times q}$, 构建向量 $z_k^x = \text{vec}(X_k)$ 和 $z_k^u = \text{vec}(U_k)$. 从一个可接受的反馈控制增益 K_k 出发, 根据系统(1)~(3)和式(15), 可得下式:

$$\begin{aligned} \dot{x} - X_k \dot{v} &= \\ Ax + Bu + Ev - X_k Sv &= \\ A(x - X_k v) + B(u - U_k v) - \\ (X_k S - AX_k - BU_k - E)v &= \\ A_k(x - X_k v) + B(u - U_k v + K_k(x - X_k v)) - \\ f_1(X_k, U_k)v. \end{aligned} \quad (19)$$

为便于描述, 针对任意非负整数 k , 本文定义矩阵 A_k 为 $A_k = A - BK_k$.

基于式(13)、(14)和(19), 可得

$$\begin{aligned} (x(t_1) - X_k v(t_1))^T P_k (x(t_1) - X_k v(t_1)) - \\ (x(t_0) - X_k v(t_0))^T P_k (x(t_0) - X_k v(t_0)) = \\ \int_{t_0}^{t_1} (x - X_k v)^T (A_k^T P_k + P_k A_k) (x - X_k v) d\tau + \\ \int_{t_0}^{t_1} 2[u - U_k v + K_k(x - X_k v)]^T B^T P_k (x - \\ X_k v) d\tau - \int_{t_0}^{t_1} 2(x - X_k v)^T P_k f_1(X_k, U_k) v d\tau = \end{aligned}$$

$$\begin{aligned}
& - \int_{t_0}^{t_1} (x - X_k v)^T (Q + K_k^T R K_k) (x - X_k v) d\tau + \\
& \int_{t_0}^{t_1} 2[u - U_k v + K_k (x - X_k v)]^T R K_{k+1} (x - \\
& X_k v) d\tau - \int_{t_0}^{t_1} 2(x - X_k v)^T P_k f_1(X_k, U_k) v d\tau. \quad (20)
\end{aligned}$$

为简化描述,定义变量 $\bar{x}_k(t) = x(t) - X_k v(t)$, $\bar{u}_k(t) = u(t) - U_k v(t)$. 通过克罗内克积的性质,有 $\bar{x}_k^T (Q + K_k^T R K_k) \bar{x}_k = (\bar{x}_k^T \otimes \bar{x}_k^T) \text{vec}(Q + K_k^T R K_k)$, $\bar{x}_k^T P_k f_1(X_k, U_k) v = (v^T \otimes \bar{x}_k^T) \text{vec}(P_k f_1(X_k, U_k))$, $(\bar{u}_k + K_k \bar{x}_k)^T R K_{k+1} \bar{x}_k = [(\bar{x}_k^T \otimes \bar{x}_k^T)(I_n \otimes K_k^T R) + (\bar{x}_k^T \otimes \bar{u}_k^T)(I_n \otimes R)] \times \text{vec}(K_{k+1})$.

对于正整数 s , 定义以下矩阵:

$$\begin{aligned}
\delta_{\bar{x}_k, \bar{x}_k} &= [\text{vecv}(\bar{x}_k(t_1)) - \text{vecv}(\bar{x}_k(t_0)), \text{vecv}(\bar{x}_k(t_2)) - \\
& \text{vecv}(\bar{x}_k(t_1)), \dots, \text{vecv}(\bar{x}_k(t_s)) - \text{vecv}(\bar{x}_k(t_{s-1}))]^T, \\
\Gamma_{\bar{x}_k, \bar{x}_k} &= \left[\int_{t_0}^{t_1} \bar{x}_k \otimes \bar{x}_k d\tau, \int_{t_1}^{t_2} \bar{x}_k \otimes \bar{x}_k d\tau, \dots, \right. \\
& \left. \int_{t_{s-1}}^{t_s} \bar{x}_k \otimes \bar{x}_k d\tau \right]^T, \\
\Gamma_{\bar{x}_k, \bar{u}_k} &= \left[\int_{t_0}^{t_1} \bar{x}_k \otimes \bar{u}_k d\tau, \int_{t_1}^{t_2} \bar{x}_k \otimes \bar{u}_k d\tau, \dots, \right. \\
& \left. \int_{t_{s-1}}^{t_s} \bar{x}_k \otimes \bar{u}_k d\tau \right]^T, \\
\Gamma_{v, \bar{x}_k} &= \left[\int_{t_0}^{t_1} v \otimes \bar{x}_k d\tau, \int_{t_1}^{t_2} v \otimes \bar{x}_k d\tau, \dots, \right. \\
& \left. \int_{t_{s-1}}^{t_s} v \otimes \bar{x}_k d\tau \right]^T,
\end{aligned}$$

其中 $t_0 < t_1 < \dots < t_s$. 基于以上矩阵定义, 可通过式(20)得到以下矩阵方程:

$$\Theta_k \begin{bmatrix} \text{vecs}(P_k) \\ \text{vec}(K_{k+1}) \\ \text{vec}(P_k f_1(X_k, U_k)) \end{bmatrix} = \Lambda_k. \quad (21)$$

其中

$$\begin{aligned}
\Theta_k &= [\delta_{\bar{x}_k, \bar{x}_k}, -2\Gamma_{\bar{x}_k, \bar{x}_k} (I_n \otimes K_k^T R) - \\
& 2\Gamma_{\bar{x}_k, \bar{u}_k} (I_n \otimes R), 2\Gamma_{v, \bar{x}_k}], \\
\Lambda_k &= -\Gamma_{\bar{x}_k, \bar{x}_k}^T \text{vec}(Q + K_k^T R K_k).
\end{aligned}$$

如果矩阵 Θ_k 列满秩, 则可通过方程(21)唯一地求解 P_k, K_{k+1} 和 $P_k f_1(X_k, U_k)$. 由于 P_k 是一个对称正定矩阵, 即非奇异矩阵, $f_1(X_k, U_k)$ 及向量 $\mathcal{R}_1(z_k^x, z_k^u)$ 均可通过方程(21)的解唯一确定. 文献[10]给出了矩阵 Θ_k 列满秩的充分条件: 如果存在正整数 s^* 使得任意的 $s > s^*$ 和时间序列 $t_0 < t_1 < \dots < t_s$, 以下

满秩条件满足

$$\text{rank}([\Gamma_{\bar{x}_k, \bar{x}_k}, \Gamma_{\bar{x}_k, \bar{u}_k}, \Gamma_{v, \bar{x}_k}]) = \frac{n(n+1)}{2} + (m+q)n, \quad (22)$$

则矩阵 Θ_k 对于任意正整数 k 都列满秩.

另根据式(16), 系统跟踪误差可描述为

$$\begin{aligned}
\int_{t_0}^{t_1} e(\tau) d\tau &= \int_{t_0}^{t_1} C x d\tau + \int_{t_0}^{t_1} F v d\tau = \\
\int_{t_0}^{t_1} C (x - X_k v) d\tau &+ \int_{t_0}^{t_1} (C X_k + F) v d\tau = \\
\int_{t_0}^{t_1} C \bar{x}_k d\tau &+ \int_{t_0}^{t_1} f_2(X_k) v d\tau = \\
\int_{t_0}^{t_1} (\bar{x}_k^T \otimes I_m) \text{vec}(C) d\tau &+ \int_{t_0}^{t_1} (v^T \otimes I_m) \mathcal{R}_2(z_k^x) d\tau. \quad (23)
\end{aligned}$$

相应地, 若定义如下矩阵:

$$\begin{aligned}
\Gamma_{\bar{x}_k} &= \left[\int_{t_0}^{t_1} \bar{x}_k \otimes I_m d\tau, \dots, \int_{t_{s-1}}^{t_s} \bar{x}_k \otimes I_m d\tau \right]^T, \\
\Gamma_e &= \left[\int_{t_0}^{t_1} e^T d\tau, \dots, \int_{t_{s-1}}^{t_s} e^T d\tau \right]^T, \\
\Gamma_v &= \left[\int_{t_0}^{t_1} v \otimes I_m d\tau, \dots, \int_{t_{s-1}}^{t_s} v \otimes I_m d\tau \right]^T,
\end{aligned}$$

则可基于在线数据学习矩阵 C 和向量 $\mathcal{R}_2(z_k^x)$, 即

$$[\Gamma_{\bar{x}_k}, \Gamma_v] \begin{bmatrix} \text{vec}(C) \\ \mathcal{R}_2(z_k^x) \end{bmatrix} = \Gamma_e. \quad (24)$$

至此, 在模型未知的情况下, 给定任意矩阵 X_k, U_k , 已经可以通过数据学习 $\mathcal{R}_1(z_k^x, z_k^u)$ 和 $\mathcal{R}_2(z_k^x)$, 其中 $z_k^x = \text{vec}(X_k)$, $z_k^u = \text{vec}(U_k)$. 如果想应用梯度下降法逼近调节器方程的解, 还需要基于数据寻找式(18)中的偏导数. 这里从 $\frac{\partial \mathcal{R}_1^T}{\partial z^x}$ 开始. 通过克罗内克积的性质, 有

$$X_k v = (v^T(t) \otimes I_n) \text{vec}(X_k) := (v^T(t) \otimes I_n) z_k^x.$$

从而可将式(20)的左边重写为

$$\begin{aligned}
\text{Left} &= (x(t_1) - X_k v(t_1))^T P_k (x(t_1) - X_k v(t_1)) - \\
& (x(t_0) - X_k v(t_0))^T P_k (x(t_0) - X_k v(t_0)) = \\
& [(x(t) - (v^T(t) \otimes I_n) z_k^x) P_k \times \\
& (x(t) - (v^T(t) \otimes I_n) z_k^x)]|_{t_0}^{t_1}. \quad (25)
\end{aligned}$$

将式(20)的右边写为

$$\begin{aligned}
\text{Right} &= \\
& - \int_{t_0}^{t_1} (x - X_k v)^T (Q + K_k^T R K_k) (x - X_k v) d\tau + \\
& \int_{t_0}^{t_1} 2[u - U_k v + K_k (x - X_k v)]^T R K_{k+1} (x - \\
& X_k v) d\tau - \int_{t_0}^{t_1} 2(x - X_k v)^T P_k f_1(X_k, U_k) v d\tau = \\
& \int_{t_0}^{t_1} (x - X_k v)^T (2K_k^T R K_{k+1} - Q - K_k^T R K_k) \times
\end{aligned}$$

$$\begin{aligned}
 & (x - X_k v) d\tau + \\
 & \int_{t_0}^{t_1} 2(u - U_k v)^T R K_{k+1} (x - X_k v) d\tau - \\
 & \int_{t_0}^{t_1} 2(x - X_k v)^T P_k f_1(X_k, U_k) v d\tau = \\
 & \int_{t_0}^{t_1} (x - (v^T \otimes I_n) z_k^x)^T (2K_k^T R K_{k+1} - \\
 & Q - K_k^T R K_k) (x - (v^T \otimes I_n) z_k^x) d\tau + \\
 & \int_{t_0}^{t_1} 2(u - U_k v)^T R K_{k+1} (x - (v^T \otimes I_n) z_k^x) d\tau - \\
 & \int_{t_0}^{t_1} 2(x - X_k v)^T P_k f_1(X_k, U_k) v d\tau. \tag{26}
 \end{aligned}$$

将式(20)的左边对 z_k^x 求偏导, 可得

$$\frac{\partial \text{Left}}{\partial z_k^x} = -2[(v(t) \otimes I_n) P_k (x(t) - X_k v(t))] \Big|_{t_0}^{t_1}. \tag{27}$$

将式(20)的右边对 z_k^x 求偏导, 可得

$$\begin{aligned}
 & \frac{\partial \text{Right}}{\partial z_k^x} = \\
 & \int_{t_0}^{t_1} -2(v \otimes I_n) (K_k^T R K_{k+1} + K_{k+1}^T R K_k - \\
 & Q - K_k^T R K_k) (x - X_k v) d\tau - \\
 & \int_{t_0}^{t_1} 2(v \otimes I_n) K_{k+1}^T R (u - U_k v) d\tau + \\
 & \int_{t_0}^{t_1} 2(v \otimes I_n) P_k f_1(X_k, U_k) v d\tau - \\
 & \frac{\partial \mathcal{R}_1^T}{\partial z_k^x} \int_{t_0}^{t_1} 2(v \otimes I_n) P_k (x - X_k v) d\tau. \tag{28}
 \end{aligned}$$

再次使用克罗内克积的性质, 有

$$\begin{aligned}
 & (v \otimes I_n) P_k (x - X_k v) = \\
 & (v \otimes I_n) (1 \otimes (P_k (x - X_k v))) = \\
 & v \otimes (P_k (x - X_k v)) = \\
 & (I_q \otimes P_k) (v \otimes (x - X_k v)). \tag{29}
 \end{aligned}$$

结合式(27)~(29), 有

$$\begin{aligned}
 & \frac{\partial \mathcal{R}_1^T}{\partial z_k^x} \int_{t_0}^{t_1} (I_q \otimes P_k) (v \otimes (x - X_k v)) d\tau = \\
 & [(I_q \otimes P_k) (v(t) \otimes (x(t) - X_k v(t)))] \Big|_{t_0}^{t_1} - \\
 & \int_{t_0}^{t_1} (v \otimes I_n) (K_k^T R K_{k+1} + K_{k+1}^T R K_k - \\
 & Q - K_k^T R K_k) (x - X_k v) d\tau - \\
 & \int_{t_0}^{t_1} (v \otimes I_n) K_{k+1}^T R (u - U_k v) d\tau + \\
 & \int_{t_0}^{t_1} (v \otimes I_n) P_k f_1(X_k, U_k) v d\tau := \\
 & \phi(t_0, t_1). \tag{30}
 \end{aligned}$$

式(30)与下式等价:

$$\left(\int_{t_0}^{t_1} (v^T \otimes (x - X_k v)^T) d\tau \right) \otimes I_{nq} \times$$

$$\text{vec} \left(\frac{\partial \mathcal{R}_1^T}{\partial z_k^x} (I_q \otimes P_k) \right) := \text{vec}(\phi(t_0, t_1)). \tag{31}$$

针对一组有限的时间序列 $t_0 < t_1 < \dots < t_s$, 定义如下矩阵:

$$\begin{aligned}
 \Psi &= \begin{bmatrix} \left(\int_{t_0}^{t_1} (v^T \otimes (x - X_k v)^T) d\tau \right) \otimes I_{nq} \\ \left(\int_{t_1}^{t_2} (v^T \otimes (x - X_k v)^T) d\tau \right) \otimes I_{nq} \\ \vdots \\ \left(\int_{t_{s-1}}^{t_s} (v^T \otimes (x - X_k v)^T) d\tau \right) \otimes I_{nq} \end{bmatrix}, \\
 \Phi &= \begin{bmatrix} \text{vec}(\phi(t_0, t_1)) \\ \text{vec}(\phi(t_1, t_2)) \\ \vdots \\ \text{vec}(\phi(t_{s-1}, t_s)) \end{bmatrix}.
 \end{aligned}$$

基于式(31)可构建一个矩阵方程

$$\Psi \text{vec} \left(\frac{\partial \mathcal{R}_1^T}{\partial z_k^x} (I_q \otimes P_k) \right) = \Phi. \tag{32}$$

通过最小二乘法, 可求得 \mathcal{R}_1^T 对 z_k^x 的偏导, 有

$$\frac{\partial \mathcal{R}_1^T}{\partial z_k^x} = \text{vec}_{nq \times nq}^{-1} ((\Psi^T \Psi)^{-1} \Psi^T \Phi) (I_q \otimes P_k)^{-1}. \tag{33}$$

假设系统矩阵全部已知, 可通过式(15)获得 $\mathcal{R}_1^T(z^x, z^u)$ 对 z^x 的偏导, 有

$$\frac{\partial \mathcal{R}_1^T}{\partial z^x} = (S \otimes I_n) - (I_q \otimes A^T). \tag{34}$$

以下引理将证明通过数据获得的偏导和通过模型获得的偏导是一致的.

引理1 当满秩条件(22)满足时, 对于任何正整数 k , 基于最小二乘法通过式(33)获得的偏导与基于系统动态信息通过式(34)获得的结果相等.

证明 通过满秩条件(22)可推出矩阵 Γ_{v, \bar{x}_k} 也是列满秩的, 这表示矩阵 Ψ 同样列满秩. 此时, 通过最小二乘法(33)可求得式(32)的唯一解. 欲证明此引理, 只要证明式(34)为(32)的一个解. 因式(32)本质上是由式(30)在不同起止时间下的方程叠加而成, 故这里只需证明对于任意的 t_0, t_1 , 将式(34)代入(30)使其成立即可.

将式(34)代入(30)的左边, 可得

$$\begin{aligned}
 & \frac{\partial \mathcal{R}_1^T}{\partial z_k^x} \int_{t_0}^{t_1} (I_q \otimes P_k) (v \otimes (x - X_k v)) d\tau = \\
 & \int_{t_0}^{t_1} (S \otimes I_n) (I_q \otimes P_k) (v \otimes (x - X_k v)) d\tau - \\
 & \int_{t_0}^{t_1} (I_q \otimes A^T) (I_q \otimes P_k) (v \otimes (x - X_k v)) d\tau = \\
 & \int_{t_0}^{t_1} [(Sv) \otimes (P_k (x - X_k v)) - v \otimes (A^T P_k (x - \\
 & X_k v))] d\tau =
 \end{aligned}$$

$$\int_{t_0}^{t_1} [\dot{v} \otimes (P_k(x - X_k v)) - v \otimes (A^T P_k(x - X_k v))] d\tau. \tag{35}$$

通过李雅普诺夫方程(13),有

$$\begin{aligned} &K_k^T R K_{k+1} + K_{k+1}^T R K_k - Q - K_k^T R K_k = \\ &K_k^T B P_k + P_k B K_k + A_k^T P_k + P_k A_k = \\ &K_k^T B P_k + P_k B K_k + (A - B_k K_k)^T P_k + \\ &P_k (A - B_k K_k) = \\ &A^T P_k + P_k A. \end{aligned}$$

式(30)的右边与下式等价:

$$\begin{aligned} &[(I_q \otimes P_k)(v(t) \otimes (x(t) - X_k v(t)))]|_{t_0}^{t_1} - \\ &\int_{t_0}^{t_1} (v \otimes I_n)(K_k^T R K_{k+1} + K_{k+1}^T R K_k - \\ &Q - K_k^T R K_k)(x - X_k v) d\tau - \\ &\int_{t_0}^{t_1} (v \otimes I_n) K_{k+1}^T R (u - U_k v) d\tau + \\ &\int_{t_0}^{t_1} (v \otimes I_n) P_k f_1(X_k, U_k) v d\tau = \\ &[v(t) \otimes (P_k(x(t) - X_k v(t)))]|_{t_0}^{t_1} - \\ &\int_{t_0}^{t_1} v \otimes (A^T P_k(x - X_k v)) d\tau - \\ &\int_{t_0}^{t_1} v \otimes (P_k A(x - X_k v)) d\tau - \\ &\int_{t_0}^{t_1} v \otimes (P_k B(u - U_k v)) d\tau + \\ &\int_{t_0}^{t_1} v \otimes (P_k f_1(X_k, U_k)) d\tau = \\ &[v(t) \otimes (P_k(x(t) - X_k v(t)))]|_{t_0}^{t_1} - \\ &\int_{t_0}^{t_1} v \otimes (A^T P_k(x - X_k v)) d\tau - \\ &\int_{t_0}^{t_1} v \otimes (P_k(\dot{x} - X_k \dot{v})) d\tau. \end{aligned} \tag{36}$$

不难看出,式(35)与(36)等价,这表示对于任意的 t_0, t_1 ,将式(34)代入(30)使其成立即可.这也表示,式(33)与(34)获得的结果等价. □

通过式(15)和(16),还可以求得另外两个偏导数,并且它们均可通过在线学习获得,即

$$\begin{aligned} \frac{\partial \mathcal{R}_1^T}{\partial z_k^u} &= -I_q \otimes B^T := -I_q \otimes (R K_{k+1} P_k^{-1}), \\ \frac{\partial \mathcal{R}_2^T}{\partial z_k^x} &= I_q \otimes C^T. \end{aligned} \tag{37}$$

至此,本文针对自适应最优输出调节问题给出一个新的在线学习算法,即算法1.

算法1 自适应最优输出调节算法.

1) 选取一个可接受的反馈增益 K_0 和任意的一

对向量 (z_0^x, z_0^u) , 令 $k = 0$.

2) 通过式(21)和(24)求得 $P_k, K_{k+1}, \mathcal{R}_1(z_k^x, z_k^u)$ 和 $\mathcal{R}_2(z_k^x)$.

3) 通过式(32)和(37)求得 $\frac{\partial \mathcal{R}_1^T}{\partial z_k^x}, \frac{\partial \mathcal{R}_1^T}{\partial z_k^u}$ 和 $\frac{\partial \mathcal{R}_2^T}{\partial z_k^x}$.

4) 通过式(18)得到 z_{k+1}^x 和 z_{k+1}^u .

5) 令 $k \leftarrow k + 1$, 重复步骤2)和步骤4)直到 $\max\{\|z_k^x - z_{k-1}^x\|, \|z_k^u - z_{k-1}^u\|, \|P_k - P_{k-1}\|\} \leq \epsilon$, 其中阈值 $\epsilon > 0$.

6) 令 $k^* \leftarrow k, X_{k^*} = \text{vec}_{n \times q}^{-1}(z_{k^*}^x), U_{k^*} = \text{vec}_{m \times q}^{-1}(z_{k^*}^u), L_{k^*} = U_{k^*} + K_{k^*} X_{k^*}$, 获得近似最优前馈-反馈控制律 $u = -K_{k^*} x + L_{k^*} v$.

文献[10]已经证明基于算法1学习获得的 P_k 和 K_k 分别收敛于黎卡提方程的解和最优反馈控制增益,即 $\lim_{k \rightarrow \infty} P_k = P^*, \lim_{k \rightarrow \infty} K_k = K^*$. 本文仅需证明 X_k 和 U_k 的收敛性. 该结果在如下定理中给出.

定理1 当满秩条件(22)满足时,通过算法1求得的解 (X_k, U_k) 收敛于调节器方程的解 (X^*, U^*) . 即 $\lim_{k \rightarrow \infty} X_k = X^*, \lim_{k \rightarrow \infty} U_k = U^*$.

证明 由引理1可知,通过式(32)求得的偏导数与基于模型求得的结果一致. 同时可以看到,通过式(37)求得的偏导数也与基于模型求得的结果一致. 另外,通过式(21)和(24)得到的向量 $\mathcal{R}_1(z_k^x, z_k^u) := \text{vec}(f_1(X_k, U_k))$ 和 $\mathcal{R}_2(z_k^x) := \text{vec}(f_2(X_k))$ 也与基于模型的结果相同. 这表明通过算法1更新得到的 $z_{k+1}^x = \text{vec}(X_{k+1})$ 和 $z_{k+1}^u = \text{vec}(U_{k+1})$ 与基于模型的、通过式(18)更新的结果是等价的. 式(18)本质上是一个梯度下降方法,当 k 趋于无穷时,迭代得到的 X_k 和 U_k 会收敛于调节器方程的解 X^* 和 U^* . □

基于定理1,经分析可以发现,算法1得到的控制器为近似最优控制器. 该控制器能够保证闭环系统的稳定性并且使闭环系统实现输出调节. 关于稳定性分析可以参考文献[10].

3 仿真实验

为验证算法1的有效性,本文针对如下线性系统进行数值仿真:

$$\begin{aligned} \dot{v}(t) &= \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} v(t), \\ \dot{x}(t) &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 0.1 \end{bmatrix} u(t), \\ e(t) &= [1 \ 0]x(t) + [-1 \ 0]v(t). \end{aligned}$$

权重矩阵分别选为 $Q = 100I_2, R = 1$. 算法1的阈值选为 $\epsilon = 0.001$. 初始控制增益选为 $K_0 = [30 \ 40]$, 同时用于梯度下降迭代的初始值为 $X =$

$0_{2 \times 2}$, $U = 0_{1 \times 2}$. 在仿真过程中, 假设所有系统矩阵均未知. 通过采集在线状态和输入数据, 利用算法1迭代矩阵 P_k 、 K_k 、 X_k 和 U_k , 其迭代结果见图1和图2. 如图所示, 当迭代次数 $k = 99$ 时, 收敛条件满足(见算法1的步骤5). 各矩阵近似值与最优值对比如下:

$$P_{99} = \begin{bmatrix} 173.205\ 08 & 99.999\ 998 \\ 99.999\ 998 & 173.205\ 08 \end{bmatrix},$$

$$P^* = \begin{bmatrix} 173.205\ 08 & 100 \\ 100 & 173.205\ 08 \end{bmatrix},$$

$$K_{99} = [9.999\ 999\ 8 \quad 17.320\ 508],$$

$$K^* = [10.000\ 0 \quad 17.320\ 508],$$

$$X_{99} = \begin{bmatrix} 0.999\ 212\ 63 & 2.959\ 623\ 7 \times 10^{-9} \\ -2.935\ 651\ 1 \times 10^{-9} & 0.998\ 685\ 36 \end{bmatrix},$$

$$X^* = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

$$U_{99} = [-9.985\ 923\ 5 \quad -0.000\ 000\ 042\ 064\ 517],$$

$$U^* = [10 \quad 0].$$

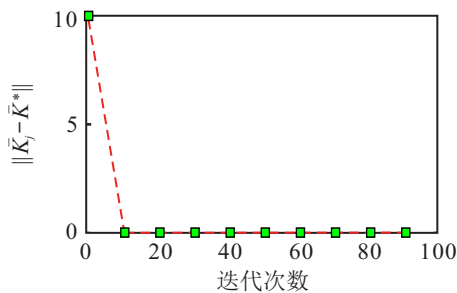
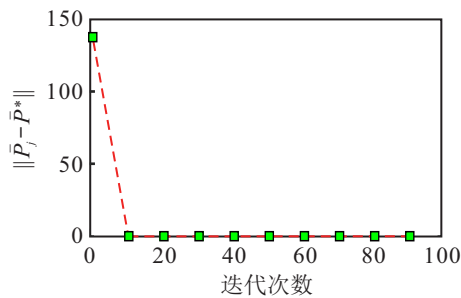


图1 应用算法1得到矩阵 (P_k, K_k) 和最优解 (P^*, K^*) 的比较

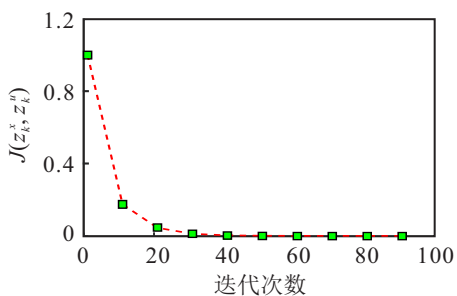


图2 应用算法1得到的成本 $J(z_k^x, z_k^u)$ (该成本越接近0, 表示迭代得到的结果与调节器方程真实解越接近)

采用学习得到的近似最优前馈-反馈控制器, 闭环系统的跟踪效果见图3. 如图3所示, 当把控制器更新为学习的控制器后, 闭环系统的输出信号可以渐近地跟踪参考信号.

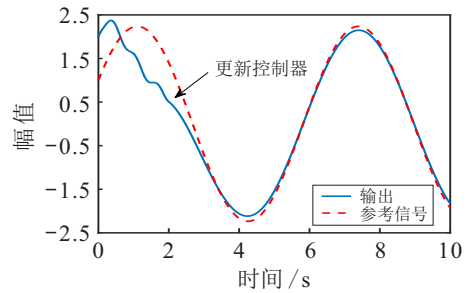


图3 应用算法1学习前后闭环系统跟踪效果比较

4 结论

本文针对连续时间动态系统的自适应最优输出调节问题, 提出了一种新的控制器设计方法. 与现有方法不同, 在系统动态未知的情况下, 该方法基于自适应动态规划和梯度下降法可同时逼近黎卡提方程和调节器方程的解析解. 未来研究将尝试针对一类非线性输出调节问题设计自适应最优控制器.

参考文献(References)

- [1] Huang J. Nonlinear output regulation: Theory and applications[M]. Philadelphia: Society for Industrial and Applied Mathematics, 2004.
- [2] 宋政一, 聂宏, 赵军. 线性离散切换系统的输出调节问题[J]. 控制与决策, 2006, 21(11): 1249-1252. (Song Z Y, Nie H, Zhao J. Output regulation of linear discrete-time switched systems[J]. Control and Decision, 2006, 21(11): 1249-1252.)
- [3] Khan G D, Chen Z Y, Zhu L J. A new approach for event-triggered stabilization and output regulation of nonlinear systems[J]. IEEE Transactions on Automatic Control, 2020, 65(8): 3592-3599.
- [4] 王澄, 魏庆来, 赵冬斌, 等. 基于数据的智能电网电能自适应优化调控[J]. 控制工程, 2014, 21(5): 753-759. (Wang C, Wei Q L, Zhao D B, et al. Data-based adaptive optimal control of smart grid power[J]. Control Engineering of China, 2014, 21(5): 753-759.)
- [5] Jiang Z P, Bian T, Gao W N. Learning-based control: A tutorial and some recent results[J]. Foundations and Trends in Systems and Control, 2020, 8(3): 176-284.
- [6] He S P, Fang H Y, Zhang M G, et al. Adaptive optimal control for a class of nonlinear systems: The online policy iteration approach[J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 31(2): 549-558.
- [7] Vamvoudakis K G, Lewis F L, Hudak G R.

- Multi-agent differential graphical games: Online adaptive learning solution for synchronization with optimality[J]. *Automatica*, 2012, 48(8): 1598-1611.
- [8] Kamalapurkar R, Walters P, Rosenfeld J, et al. Model-based reinforcement learning for approximate optimal control[C]. *Reinforcement Learning for Optimal Feedback Control*. Cham: Springer, 2018: 99-148.
- [9] Yang Y L, Modares H, Vamvoudakis K G, et al. Hamiltonian-driven adaptive dynamic programming with approximation errors[J]. *IEEE Transactions on Cybernetics*, 2022, 52(12): 13762-13773.
- [10] Gao W N, Jiang Z P. Adaptive dynamic programming and adaptive optimal output regulation of linear systems[J]. *IEEE Transactions on Automatic Control*, 2016, 61(12): 4164-4169.
- [11] Gao W N, Jiang Z P. Learning-based adaptive optimal output regulation of linear and nonlinear systems: An overview[J]. *Control Theory and Technology*, 2022, 20(1): 1-19.
- [12] Xie K D, Yu X, Lan W Y. Optimal output regulation for unknown continuous-time linear systems by internal model and adaptive dynamic programming[J]. *Automatica*, 2022, 146: 110564.
- [13] Wu Y Z, Liang Q P, Hu J P. Optimal output regulation for general linear systems via adaptive dynamic programming[J]. *IEEE Transactions on Cybernetics*, 2022, 52(11): 11916-11926.
- [14] Chen C, Modares H, Xie K, et al. Reinforcement learning-based adaptive optimal exponential tracking control of linear systems with unknown dynamics[J]. *IEEE Transactions on Automatic Control*, 2019, 64(11): 4423-4438.
- [15] Kleinman D. On an iterative technique for Riccati equation computations[J]. *IEEE Transactions on Automatic Control*, 1968, 13(1): 114-115.

作者简介

高伟男(1990—),男,教授,博士生导师,从事人工智能、自适应动态规划、优化控制、智能网联汽车等研究, E-mail: gaown@mail.neu.edu.cn;

杨涛(1981—),男,教授,博士生导师,从事信息物理系统、复杂工业过程的分布式优化等研究, E-mail: yangtao@mail.neu.edu.cn;

柴天佑(1947—),男,教授,博士生导师,中国工程院院士,从事自适应控制、智能解耦控制、流程工业综合自动化与智能化系统理论、方法与技术等研究, E-mail: tychai@mail.neu.edu.cn.



特邀专家 高伟男,东北大学流程工业综合自动化国家重点实验室教授、博士生导师,IEEE高级会员,美国波士顿三菱电气研究实验室客座教授. 2011年和2013年分获东北大学学士和硕士学位,2017年获美国纽约大学博士学位. 曾任美国佛罗里达理工学院助理教授、博士生导师. 长期从事人工智能、自适应动态规划、优化控制和输出调节等基础理论研究,并以智能网联汽车、无人驾驶、电力系统等为具体对象进行基础应用研究. 任《自动化学报(英文版)》“IEEE TNNLS”“Control Engineering Practice”等控制与人工智能领域国际期刊编委,美国控制会议程序委员会委员,中国工业人工智能专委会委员. 获IEEE数据驱动控制与学习系统会议最佳论文奖、IEEE实时计算与机器人国际会议控制最佳论文奖、纽约大学David Goodman研究奖.

专家寄语 百年育人,桃李成林,世纪校庆,师生同心. 值此百年校庆之际,衷心祝愿母校宏图更展,永远闪耀璀璨光芒!