

基于鲁棒最优控制的在线机器学习方法

宁瀚文, 李占风*

中南财经政法大学统计与数学学院, 武汉 430073

E-mail: ninghanwen@gmail.com, lizf63@aliyun.com

收稿日期: 2017-05-16; 接受日期: 2017-10-18; 网络出版日期: 2018-06-04; * 通信作者

国家自然科学基金 (批准号: 11301544 和 61773401) 和国家留学基金委 (批准号: 201707085011) 资助项目

摘要 基于再生核 Hilbert 空间 (reproducing kernel Hilbert space, RKHS) 的统计学习模型被广泛应用于函数逼近、图像处理、模式识别和回归分析等领域, 并且也在非线性随机动力学系统的辨识问题中有着很好的表现. 本文提出一个基于鲁棒最优控制的 RKHS 模型学习方法, 来实现对非线性随机动力学系统的高效在线建模. 利用本文得到的关于再生核空间的一些理论结果, 本文将随机动力学系统的在线学习问题转化为一组具有有界随机扰动的离散时变线性系统的输出反馈镇定问题, 并利用模型预测控制技术来设计相应的控制算法和学习算法. 与现有的 RKHS 模型学习方法相比, 在不引入任何数据窗口原理、剪枝技术、学习步长的调整机制以及对噪声统计性质的假设的情形下, 新方法可以在保证模型参数快速且鲁棒收敛的同时, 实现对动力学系统的自适应高精度建模. 此外, 本文首次从最优控制的视角出发, 研究动力学系统的在线核学习问题. 在本文提出的研究框架下, 现有各种控制技术可以被利用起来开发新的鲁棒学习方法, 这也为核学习理论的研究和算法的开发提供一些新的思路. 本文亦给出了数值算例和对比结果, 用来说明新方法的有效性.

关键词 统计学习 在线学习 再生核空间 随机动力学系统 鲁棒最优控制 模型预测控制

MSC (2010) 主题分类 62J02, 68T05, 93A30, 93E35

1 引言

随机动力学系统的常规模型包括随机普通微分方程^[1]、随机偏微分方程^[2,3]和非线性自回归模型^[4]、子空间模型^[5]和 Wiener 系统^[6,7]等. 这些系统的统计学习模型是系统状态控制、分析和设计的基础^[8-10]. 统计学习模型不仅可以用于理解系统的动力学特征, 还可以用来预测系统未来的行为. 现有统计学习方法包括集合成员^[11]、神经网络^[12,13]和正交最小二乘^[14,15]等. 在现有的诸多方法中, 基于再生核 Hilbert 空间理论的学习方法成为最流行的方法之一^[16-18]. 一些强大的核学习算法, 如支持向量机^[19,20]、多核方法^[21-23]和简化核方法^[24]等, 被研究出来并应用于动力学系统的建模中.

英文引用格式: Ning H W, Li Z F. An adaptive online machine learning method based on a robust optimal control approach (in Chinese). *Sci Sin Math*, 2018, 48: 1181-1202, doi: 10.1360/SCM-2017-0308

在许多实际的建模任务中, 时间序列数据不断到达, 统计学习模型需要提供及时的响应, 如网络点击数据、股票交易数据和机械损伤反馈数据等. 因此, 动力学系统的在线学习方法也受到越来越多的关注^[25]. 在在线学习问题中, 如果系统特征是时变的, 则估计模型必须实时更新以追踪动力学系统的变化. 一般来讲, 在线学习有两大要求, 其一是在获得新的观测数据时, 逐步更新模型, 而不是从头开始训练样本. 其二则是在学习过程中去除旧的和无用的样本和信息, 以保持系统的简约性, 并将计算复杂度保持在可接受的水平. 基于 RKHS 的模型主要存在两类在线学习方法. 一类是在线批量学习方法, 如递归支持向量机和移动窗口最小二乘法等^[26,27]. 在这些方法中, 新的样本不断地进入到学习中, 经验误差指标也根据样本的变化进行相应的修改, 以用于更新估计模型. 修剪误差最小原则^[28]、样本字典策略^[29]和样本选择原则^[30]等剪枝技术被用来去掉无用样本, 以避免训练时间随样本数量的膨胀而非线性地增加. 另一类方法是增量优化学习方法^[31,32], 其中最具代表性的是梯度算法. 在这些方法中, 梯度下降准则被引入具有正则化项的经验误差指标的优化中, 以得到最优解及模型参数的更新算法. 在这些算法中, 计算量不会随着数据量的增加而过快增加. 一些文献也对这类方法的收敛速度和逼近能力进行了讨论^[33,34].

尽管各种在线学习方法的有效性已经被它们在实际应用领域的巨大作用所证实, 但是在遇到复杂的数据流时, 现有的方法仍然存在一些局限性. 一方面, 在线批量学习方法大多是基于标准最小二乘形式的经验误差函数, 这潜在地假设了系统噪声是白噪声过程. 这个假设显然在很多情形并不适用, 因为许多实际的数据流也经常受到具有时变特征的噪声的干扰. 如果应用现有的方法将这些动力学过程不假思索地当成白噪声扰动的数据流进行处理, 那么相应算法的学习速度、鲁棒性和精度都会受到较大程度的影响^[17,35,36]. 一些文献针对自相关、非 Gauss 或异方差的噪声扰动系统提出了一些改进的统计学习方法, 但是它们仍然不能够被轻易地推广到在线学习中. 这是因为对于时变动力学系统而言, 获得稳定的噪声项的特征参数估计往往比较困难, 并且计算效率较低^[35,36]. 另一方面, 在增量优化在线学习方法中, 一些重要的系数, 例如, 学习步长必须适应性地满足一些与动力学性质相关的特定的数值约束, 以保证梯度算法的收敛性, 因而, 最优的系数需要随着动力学系统的变化而进行调整^[37,38]. 然而, 现有的大多数在线核学习方法都没有提供这样的调整机制, 这就导致在一些实际应用中这些方法可能存在一定的局限性.

利用本文建立的一些新的理论结果, 我们将随机动力学系统的在线学习问题合理地转化为一组具有有界随机扰动的离散时变线性系统的输出反馈镇定问题, 并利用模型预测控制技术^[39-46]来设计相应的控制算法和学习算法. 与现有的 RKHS 学习方法相比, 在不引入任何数据窗口原理、剪枝技术、学习步长以及对噪声统计性质的假设的情形下, 新方法可以在保证模型参数快速收敛的同时, 得到高精度的、自适应的和更为鲁棒的建模效果. 此外, 本文首次试图从最优控制的角度来研究动力学系统的学习问题, 在本文提出的学习框架下, 许多经典的控制技术可以被借鉴过来解决诸如在线分类、图像处理和模式识别等重要的学习问题, 因而也为核学习理论研究和算法的开发提供一些新的思路.

本文的结构如下. 第 2 节给出了一些基本的定义并对现有的在线核学习算法进行了简单的梳理. 第 3 节提出了自适应鲁棒在线学习方法. 第 4 节对新方法进行了一些必要的讨论和分析. 第 5 节提供了数值算例来验证新方法的有效性. 在最后一节中, 我们对本文进行了总结.

2 现有的在线核学习方法

本节简要回顾 RKHS 的基本概念和一些经典的在线核学习方法.

再生核 Hilbert 空间构成了学习理论中一类重要的函数空间. 记 H_K 为与核 K 和内积 $\langle \cdot, \cdot \rangle_H$ 相关的再生核 Hilbert 空间, 这使得 K 具有以下再生性^[47]:

$$(1) \langle f, K(x, \cdot) \rangle_H = f(x), \langle K(x_1, \cdot), K(x_2, \cdot) \rangle_H = K(x_1, x_2), \text{ 对 } \forall x, x_1, x_2 \in X, f \in H_K.$$

(2) H_K 是所有 $K(x, \cdot)$ 张成的闭包, 这意味着 $f \in H_K$ 是 x 取不同值时核函数 $K(x, \cdot)$ 的线性组合.

在 $f \in H_K$ 上, 内积记为 $\|f\|_H = \langle f, f \rangle_H^{\frac{1}{2}}$. 通常有 $X \subset \mathbb{R}^d$, 常用的核函数包括多项式核 $K_{\text{pol}}(x_i, x_j) = (x_i^T x_j + r)^{d_0}$ 和径向基函数 (RBF) $K_{\text{RBF}}(x_i, x_j) = \exp(-\|x_i - x_j\|_2^2 / \sigma^2)$, 其中 r, d_0 和 σ 为调节参数. 在后续的章节中, 假设 X 是 \mathbb{R}^d 中的一个紧集.

2.1 批量在线学习方法

给定数据流 $(x(t), y(t))$ ($t = 1, 2, \dots$) 且假设数据流由如下模型所生成, 即 $y(t) = g^*(x(t)) + e(t)$, $e(t)$ 为零均值随机噪声, g^* 可能是时变的. 在线学习即根据数据流计算 g 以估计 g^* , 并且实时更新 g 以追踪 g^* 的变化. 为了实现这一目标, 假设在学习过程一开始, 我们有数据集 $(x(t), y(t))$, $t = 1, 2, \dots, N$, 并设计优化问题^[17, 19]

$$\min_{g \in \mathcal{H}_K} \sum_{t=1}^N (y(t) - g(x(t)))^2 + \gamma \|g\|_H^2. \quad (2.1)$$

在 H_K 中进行优化, 可以得到估计模型

$$\hat{y}(t) = \sum_{k=1}^N \alpha_k K(x(k), x(t)), \quad (2.2)$$

模型参数可以通过求解线性矩阵方程 $[\Omega + 1/\gamma I]\alpha = Y$ 得到, $Y = (y(1), \dots, y(N))^T$, $\Omega = [K_{i,j}]_{i,j=1,\dots,N}$ $= K(x(i), x(j))$. 上述批量学习方法不能直接应用于在线情形, 因为在学习过程中所有的样本都将被逐步引入到模型中. 随着矩阵 Ω 规模的增加, 会逐渐导致计算上不可行, 所估计的模型将缺乏稀疏性. 因此需要引入样本选择策略或剪枝技术来解决稀疏问题. 文献 [48] 提出了移动窗口方法, 它假定最先进来的一个样本包含最少的有效信息, 因而在建模过程中逐步删除那些先进入学习的样本. 在文献 [48] 中, 对应最小 Lagrange 乘数的样本被认为给学习模型带来的逼近能力是最小的, 因而被认为是最无用的样本. 另一种流行的样本修剪方法称为修剪误差最小化方法或快速删除方法, 即选择删除后使得模型逼近误差最小的样本^[26, 28]. 通过修剪, 训练集的大小不会大幅度增加, 因而保证了学习过程在计算上的可行性. 值得注意的是, 在这些方法中, 稀疏解一般仍然依赖于求解一个标准二次优化问题, 如 (2.1). 因而, 潜在地对动力学系统的随机部分做出标准正态性假设. 尽管许多重要的建模问题可以在这样的假设中得到解决, 但是现实的数据经常受到异方差、序列相关, 甚至是非 Gauss 非稳定噪声的扰动. 在这些情形下, 使用标准二次优化指标来进行参数估计的话, 学习效果可能不尽如人意^[35, 36]. 学者们也提出了一些改进方法. 例如, 通常利用加权最小二乘法来处理具有异质性的随机扰动的动力学系统:

$$\min_{g \in \mathcal{H}_K} \sum \sigma_t e(t)^2 + \gamma \|g\|_H^2, \quad \text{s.t. } e(t) = y(t) - g(x(t)), \quad (2.3)$$

其中 σ_t 为权重. 第一步, 检验异方差的存在性. 第二步, 估计随机过程的方差变化以得到 σ_t . 第三步, 使用第二步得到的 σ_t , 求解加权最小二乘优化问题 (2.3). 对于序列相关的噪声扰动过程的学习, 也需要

采取类似的步骤. 首先, 对时间序列数据的自相关性进行统计检验. 其次, 估计自相关阶数和对应系数. 最后改进优化问题并求解以获得更好的参数估计. 进一步地, 考虑定义在时域 $[t_0, t_1] \cup (t_1, t_2] \cup (t_2, t_3]$ 上的系统 $y(t) = f(x(t)) + e(t)$. 假设 $y(t) = f_1(x(t)) + e_1(t)$, $t \in [t_0, t_1]$; $y(t) = f_2(x(t)) + e_2(t)$, $t \in [t_1, t_2]$; $y(t) = f_3(x(t)) + e_3(t)$, $t \in [t_2, t_3]$, 其中 $e_1(t)$ 为白噪声, $e_2(t)$ 是自相关过程, $e_3(t)$ 是正态分布但具有时变的方差, f_1 、 f_2 和 f_3 是完全不同的函数且快速切换. 显然, 现有的在线学习方法在处理这种情形时会变得非常复杂甚至是不可行的. 因此, 开发自适应的鲁棒在线学习方法是非常有必要的. 而且据我们所知, 很少有文献涉及这一问题.

2.2 增量优化在线学习方法

核方法与梯度下降算法的结合为基于再生核 Hilbert 空间模型的估计提供了一个有效的研究路径 [17, 29, 31, 38, 49]. 令 $(x(t), y(t))$ ($t = 1, 2, \dots$) 为一个样本数据流, 考虑增量优化学习方法, 在 $t + 1$ 时刻, f 的估计函数 f_t 可以更新为

$$f_{t+1} = f_t + \eta_t[(y(t) - f_t(x(t)))K(x(t), \cdot) + \lambda f_t], \quad (2.4)$$

其中 $\lambda > 0$ 是正则化参数, η_t 是学习步长. 通常设 f_1 为 0. 也可以通过求解不含正则项的优化指标来进行参数估计与模型的更新, 即

$$f_{t+1} = f_t + \eta_t(y(t) - f_t(x(t)))K(x(t), \cdot). \quad (2.5)$$

上述方法也称为随机逼近算法, 更多的细节参见文献 [50–52]. 众所周知, 学习步长在算法中起着至关重要的作用. 一个较大的学习步长可以使我们快速地学习到系统的动力学性质, 但也有导致算法不稳定的风险. 一个很小的学习步长可以保证渐近收敛, 但也可能导致学习速度过慢, 使得算法在对快速变化的动力学系统的建模中适用性和预测精度较差, 因而, 最好的解决方案并不是通过尽可能小的学习步长获得的. 大多数的基于梯度的方法在很大程度上都依赖于学习步长的选择, 但是问题的复杂性在于, 就单个问题而言, 一个好的学习步长的范围通常较小, 而不同问题的最优的学习步长往往相差甚远, 换言之, 在特定问题中, 即使是经验丰富的实验者也经常不得不手动相应的最优值 [32]. 但是, 显然最优的学习率不可能完全根据经验确定. 此外, 为了保证算法的稳定性, η_t 通常需要满足一些与样本性质和正则化条件相关的约束. 例如, 在文献 [31, 51, 52] 中, 要求 $\eta_t < N_0/\text{tr}[R]$, 其中 N_0 是学习所用到的样本个数, R 是对应的核 Gram 矩阵. 在文献 [53] 中, 步长需满足 $\eta_t < \lambda/C_{l,K}$, 其中 $C_{l,K}$ 是时变且依赖于样本的常数. 另外, 在在线学习问题中, 目标函数往往是快速变化的, 这意味着学习步长的最优值也应该是时变的. 因此, 从收敛的角度来看, 最好的解决方案也无法通过经验的方式来获得. 虽然一些文献给出了一些具有适应性步长策略的方法来解决梯度算法中的局部最小值和学习速度问题 [54, 55], 然而, 这些算法通常需要启发式知识或额外的成本来调整学习步长, 难以保证所得到的在线学习算法的收敛性和鲁棒性, 也没有考虑本文所讨论的噪声效应问题.

3 鲁棒自适应在线学习方法

3.1 误差动力学系统

本小节试图从最优控制的角度提出新的思路和算法来解决本文前面所提到的问题. 假设有通过 $y(n) = f^*(x(n)) - \varepsilon(n)$ 生成的数据流 $(x(1), y(1)), (x(2), y(2)), \dots, (x(n), y(n)), \dots$, 其中 f^* 和 ε 分别为

未知非线性函数和零均值的随机噪声过程, 对于任意 f^* , 设在再生核空间 H_K 中存在一个理想的 w^* , 使得

$$y(n) = \phi(x(n))w^* - \varepsilon(n), \quad (3.1)$$

并且在每一个离散的时刻 n , 有如下估计模型:

$$\hat{y}(n) = \phi(x(n))w(n), \quad (3.2)$$

其中 ϕ 是从状态空间 \mathbb{R}^d 到 RKHS 的特征映射, 并且 $(\phi(x(i)), \phi(x(j))) = K(x(i), x(j))$. 假设 $(x(1), x(2), \dots) \in X$, X 为 \mathbb{R}^m 的一个紧集. n 时刻的估计误差为 $e(n)$, 由下式给出:

$$e(n) = \hat{y}(n) - y(n) = \phi(x(n))(w(n) - w^*) + \varepsilon(n) = \phi(x(n))w(n) - y(n). \quad (3.3)$$

因此有

$$e(n+1) - e(n) = \phi(x(n+1))w(n+1) - \phi(x(n))w(n) - y(n+1) + y(n). \quad (3.4)$$

令 $\Delta\phi(x(n+1)) = \phi(x(n+1)) - \phi(x(n))$, $\Delta w(n) = w(n+1) - w(n)$ 和 $\Delta y(n) = y(n+1) - y(n)$, 得到

$$e(n+1) - e(n) = (\phi(x(n)) + \Delta\phi(x(n)))(w(n) + \Delta w(n)) - \phi(x(n))w(n) - y(n+1) + y(n), \quad (3.5)$$

这意味着

$$e(n+1) = e(n) + \phi(x(n+1))\Delta w(n) + d^*(n), \quad (3.6)$$

其中 $d^*(n) = \Delta\phi(x(n))(w(n) - w^*) + \varepsilon(n) - \varepsilon(n+1)$ 也可以被看作一个随机扰动. 我们注意到, 如果没有仔细的设计, 这个误差系统可能不会被传统的参数更新算法所稳定. 例如, 对于经典的核最小均方 (kernel least mean square, KLMS) 算法^[31], 学习策略可以写成 $\Delta w(n) = -\eta\phi(x(n))e(n)$, 其中 $\eta > 0$ 是学习步长. 将其代入 (3.6), 我们有 $e(n+1) = [1 - \eta K(x(n), x(n+1))]e(n) + d^*(n)$. 由于 $1 - \eta K(x(n), x(n+1))$ 是时变的, 那么当学习步长固定时其并不总在 $(-1, 1)$ 的范围内, 所以, KLMS 可能无法保证误差系统的收敛性. 此外, 从泛化和正则化的角度来看, 最好的解也不是以设置尽可能小的步长来获得的^[31]. 为了解决这些问题, 一些改进策略应运而生, $w(n) = -\eta\phi(x(n))e(n) + a\Delta w(n-1)$, 即 $w(n) = -\eta\phi(x(n))e(n)/(1 - a\mathcal{B})$ ^[54], 其中 $0 < a < 1$ 且 \mathcal{B} 是滞后算子, 将这个更新代入 (3.6), 有

$$e(n+1) = [1 + a - \eta K(x(n), x(n+1))]e(n) - ae(n-1) + d^*(n) - ad^*(n-1). \quad (3.7)$$

显然, 这是一个带时变参数的二阶自回归系统, 系统的收敛性仍然无法得到保证. 注意到 $w(n+1) - w^* = w(n) - w^* + \Delta w(n)$, 则 (3.6) 可以写成

$$\begin{bmatrix} e(n+1) \\ w(n+1) - w^* \end{bmatrix} = \begin{bmatrix} 1 & \Delta\phi(x(n)) \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} e(n) \\ w(n) - w^* \end{bmatrix} + \begin{bmatrix} \phi(x(n+1)) \\ 1 \end{bmatrix} \Delta w(n) + \begin{bmatrix} d(n) \\ 0 \end{bmatrix}, \quad (3.8)$$

其中 $d(n) = \varepsilon(n) - \varepsilon(n+1)$, 那么误差动力学系统 (3.8) 可以被视作以预测误差 $e(n)$ 和参数估计误差 $w(n) - w^*$ 为状态变量、 $\phi(x(n+1))$ 为系统参数和 $\Delta w(n)$ 为控制输入的无穷维离散线性时变动力学系统. 在现有文献中, 往往假设 $\varepsilon(n)$ 和 $y(t)$ 服从正态分布, 以便对系统进行统计分析^[10, 20, 56]. 然而,

我们通常研究的系统一般都有着有界性的特征 (如在 GARCH 模型中, 收益率虽然被设定为正态分布或者学生 t 分布, 但是事实上它们是有界的). 因而, 在本文中, 为了更好地阐述我们的思想, 假定 $\varepsilon(n)$ 和 $d(n)$ 均被限制在一个紧集 $(-D, D)$ 内. 考虑到向量 $w(n)$ 一般是无穷维 Hilbert 空间中的系数, 为了将 (3.8) 转化为一个有限维系统的学习问题, 我们希望在 H_K 中的一个足够稠密的有限维子空间中对目标函数进行逼近^[24, 27]. 给定 $u_i \in X$ ($i = 1, 2, \dots, M$), 其中 u_i 是不同的向量, 令 H_K^s 为由基向量 $\phi(u_i)$ ($i = 1, \dots, M$) 张成的 H_K 的线性子空间. 那么, $w(n)$ 和 w^* 可以写成

$$w(n) = \sum_{i=1}^M \alpha_i(n) \phi(u_i), \quad w^* = \sum_{i=1}^M \alpha_i^* \phi(u_i), \quad (3.9)$$

其中 α_i 和 α^* 是 H_K^s 中的系数. 记 $\alpha(n) = (\alpha_1(n), \alpha_2(n), \dots, \alpha_M(n))$, $\alpha^* = (\alpha_1^*, \alpha_2^*, \dots, \alpha_M^*)$ 以及 $\overline{\alpha_i(n)} = \alpha_i(n) - \alpha_i^*$, $\Delta\alpha_i(n) = \alpha_i(n+1) - \alpha_i(n)$, $\Delta K(u_i, x(n)) = \Delta K(u_i, x(n+1)) - \Delta K(u_i, x(n))$, $i = 1, 2, \dots, M$. 因此有

$$e(n+1) = e(n) + \sum_{i=1}^M \Delta\alpha_i(n) K(u_i, x(n+1)) + \sum_{i=1}^M \overline{\alpha_i(n)} \Delta K(u_i, x(n)) + d(n), \quad (3.10)$$

其中 $\overline{\alpha(n)} = (\overline{\alpha_1(n)}, \dots, \overline{\alpha_M(n)}) = (\alpha_1(n) - \alpha_1^*, \alpha_2(n) - \alpha_2^*, \dots, \alpha_M(n) - \alpha_M^*)$, $\bar{K}(n) = (K(u_1, x(n)), \dots, K(u_M, x(n)))^T$, $\Delta\alpha(n) = (\Delta\alpha_1(n), \dots, \Delta\alpha_M(n))$. (3.10) 可以写为

$$e(n+1) = e(n) + \overline{\alpha(n)} \Delta \bar{K}(n) + \bar{K}(n+1) \Delta\alpha(n) + d(n). \quad (3.11)$$

注意到 $\overline{\alpha(n+1)} = \alpha(n+1) - \alpha^* = \alpha(n) - \alpha^* + \Delta\alpha(n) = \overline{\alpha(n)} + \Delta\alpha(n)$, 令 $I_{M,M}$ 为 $M \times M$ 维的单位阵, $0_{M,1}$ 为 $M \times 1$ 维的向量, 则无穷维误差系统 (3.8) 可以被如下的有限维系统来近似:

$$E(n+1) = A(n)E(n) + B(n)\Delta\alpha(n) + Ld(n), \quad (3.12)$$

其中

$$E(n) = (e(n), \overline{\alpha(n)})^T, \quad A(n) = \begin{bmatrix} 1 & \Delta \bar{K}(n) \\ 0_{M,1} & I_{M,M} \end{bmatrix}, \quad B(n) = (\bar{K}(n+1), I_{M,M})^T, \quad L = (1, 0_{M,1}^T)^T.$$

(3.12) 可以被视为一个带有有界扰动和控制输入 $\Delta\alpha(n)$ 的离散线性系统. 需要指出的是, 在 n 时刻, 虽然 (3.12) 中所涉及的 $x(n+1)$ 是第 $n+1$ 时刻观测值, 这依然是合理的. 例如, 在非线性自回归模型 $y(n+1) = f(y(n), y(n-1)) + e(n+1)$ 和非线性随机常微分方程 $dy(t) = f(y(t))dt + g(y(t))dW(t)$ 中, n 时刻系统的状态输出是第 $n+1$ 时刻的统计学习模型的输入变量之一^[4, 21, 57]. 如果 $x(n+1)$ 在 n 时刻未知, 我们也可以将学习过程滞后一期以应用我们的模型.

通过 (3.12), 在线学习问题被转化为一个具有有限维输入控制 $\Delta\alpha(n)$ 的离散线性时变系统的镇定 (反馈控制) 问题, 就是设计最优的控制输入 $\Delta\alpha(n)$, 亦即参数的更新策略使得辨识误差 $e(n)$ 和参数的估计误差 $\alpha(n) - \alpha^*$ 所构成的系统 (3.12) 的状态变量尽可能收敛到原点. 为了进行在线学习, 要对 $\Delta\alpha(n)$ 根据预测误差进行计算以响应观测值. 注意到 $\overline{\alpha(n)} = \alpha(n) - \alpha^*$ 无法观测, 仅有误差项 $e(n)$ 可以被观测, 我们需要设计输出反馈控制策略 $\Delta\alpha(n) = F(n)CE(n) = F(n)e(n)$ 来稳定 (3.12), 其中 $C = (1, 0_{M,1}^T)$. 尽管系统中随机扰动一直存在, 使得完全的渐近收敛性 (即 $\lim_{n \rightarrow \infty} E(n) = 0$) 不可能

达到, 然而, 一个适合的控制输入可以使得误差系统收敛到一个较小紧集内. 假定进行输出反馈控制之后的闭环系统为

$$E(n+1) = (A(n) + B(n)F(n)C)E(n) + Ld(n). \quad (3.13)$$

用 $H(n) = A(n) + B(n)F(n)C$ 来表示闭环的状态转移矩阵. 对 $e(n+k)$, 我们有

$$E(n+k) = \left(\prod_{l=0}^k H(n+l) \right) E(n) + \sum_{l=1}^k \left(\prod_{j=1}^l H(n+j) \right) Ld(n+j-1). \quad (3.14)$$

注意到 $|d(n)| < D < \infty$, 令 $\lambda_{\max} H$ 表示 H 特征值模的上确界. 如果对于任意的 n , $H(n)$ 是压缩的, 即存在常数 h_* 使得 $\lambda_{\max} H(n) < h_* < 1$, 则有

$$\|E(n+k)\| < h_*^k \|E(n)\| + \sum_{l=1}^k e_*^l D = h_*^k e(n) + D \frac{1 - h_*^{k+1}}{1 - h_*}. \quad (3.15)$$

可以看到, 在适当的控制输入 $\Delta\alpha(n)$ 下, 当 $k \rightarrow \infty$ 时, $e(n)$ 和 $\alpha(n) - \alpha^*$ 会收敛到一个紧集中, 从而意味着我们可以通过对 $\Delta\alpha(n)$ 适当地设计得到在线学习算法和模型参数的估计. 此外, 如果 h_* 很小, 那么参数 w^* 的估计误差也会足够小, $e(n)$ 的大小也不会超过噪声的上确界.

这里给出一些关于 RKHS 的性质的理论结果来进一步说明上述方法的合理性. 考虑子空间 H_K^s , 给定紧集 $B \subset \mathbb{R}^d$ 和基向量 $\phi(u_i)$, $i = 1, \dots, M$, 其中 $u_i \in B$, 对任意 $u \in B$, $\phi(u) \in H_K$ 用 $\phi(u_i)$ 的线性组合来近似, 即

$$\phi(u) = \sum_{i=1}^M c(i) \phi(u_i) + \phi_{\text{res}}, \quad (3.16)$$

其中 ϕ_{res} 表示残差向量, $c = \{c(i)\}_{i=1, \dots, M}$ 由如下 RKHS 中的线性回归得到:

$$c = \min_c \left\| \sum_{i=1}^M c(i) \phi(u_i) - \phi(u) \right\|^2 = \min_c \|\phi_{\text{res}}\|^2. \quad (3.17)$$

该回归也称为近似线性依赖 (approximate linear dependent, ALD). 残差分量的大小可以衡量 $\phi(u)$ 由基向量表示的程度. 对于特征映射 ϕ 和 H_K^s 中的系数, 我们建立以下引理:

引理 3.1 特征映射 ϕ 是一个紧映射, 即 ϕ 可以把 \mathbb{R}^d 中的有界集映射到 RKHS 中的相对紧集. 此外, 对于给定的紧集 $B \in \mathbb{R}^d$ 和 $\forall \delta > 0$, 存在有限的 M 个基向量 $\phi(u_i)$, 其中 $u_i \in B$, $i = 1, 2, \dots, M$, 使得对于任意的 $\phi(u) \in H_K$, $u \in B$, 残差向量的模满足如下近似线性相关条件:

$$\|\phi_{\text{res}}\|^2 = \min_c \left\| \sum_{j=1}^m c(j) \phi(u_j) - \phi(u) \right\|^2 < \delta. \quad (3.18)$$

引理 3.2 假设 $i \neq j$, $u_i \neq u_j$. 如果不存在不全为 0 的向量 (b_1, b_2, \dots, b_M) , 使得 $\sum_{j=1}^M b_j \phi(u_j) = 0$, 那么, 核矩阵 $[K(u_i, u_j)]_{1 \leq i, j \leq M}$ 是正定矩阵. 对任意 $w \in H_K^s$, 存在唯一的 c , 使得

$$w = \sum_{i=1}^M c(i) \phi(u_i),$$

这意味着如果存在 c_1 和 c_2 , 使得 $w = \sum_{i=1}^M c_1(i) \phi(u_i) = \sum_{i=1}^M c_2(i) \phi(u_i)$, 必有 $c_1 = c_2$.

上述引理说明了, 如果合理选择了 u_i (后文将指出 u_i 的选择方法), $\|\phi_{\text{res}}\|$ 可以达到充分小. 这意味着, 如果选择适当的基向量, 在给定的紧集上, 我们在有限维空间 H_K^s 中求取的最优解与在 H_K 中求取的最优解, 它们的差距会非常小. 可以看出 H_K^s 与 \mathbb{R}^M 是同构的, 这意涵着, 在 H_K 空间中的无穷维误差动力学系统 (3.8) 与有限维误差动力学系统 (3.12) 是一一对应的, 因此, 用 (3.12) 来近似地表达 (3.8), 并利用 (3.12) 来设计学习算法是完全合理的. 这两个引理进一步说明了我们方法的合理性.

3.2 模型预测控制与在线自适应学习算法设计

模型预测控制 (model predictive control, MPC) [58, 59] 是一种可靠的鲁棒最优控制策略, 它可以用于高性能模型的过程控制系统的设计. MPC 也可以被认为是基于在线优化的方法, 即可以利用受控系统模型来预测系统从当前状态开始, 沿着给定预测时间范围的状态演变. 在 MPC 控制中, 基本思想是通过对一个系统状态和控制输入相关的二次型性能指标优化, 得到一组给定未来时域上的控制输入来对系统响应进行控制与预测. 一旦优化问题被计算解决, 这组时域上的控制序列中仅有第一个控制输入被使用, 其余的则被忽略. 这个优化过程在下一个采样步骤中重复, 从而逐步得到反馈控制镇定受控系统. MPC 也称为后退时域控制. 本文利用 MPC 技术, 来设计系统 (3.12) 输出反馈控制率, 从而建立有效的参数更新策略和核学习方法. 对于任意时刻 n , 我们构造基于 (3.12) 的虚拟时不变系统

$$E_n(t+1) = A_n E_n(t) + B_n \Delta \alpha_n(t) + L d_n(t), \quad t = 1, 2, \dots, \quad (3.19)$$

其中 $A_n = A(n)$, $B_n = B(n)$ 与 (3.12) 中的定义相同. $\Delta \alpha_n(t) = \alpha_n(t+1) - \alpha_n(t)$ 为输出反馈控制序列, 且 $\Delta \alpha_n(1) = \Delta \alpha(n) = F_n C E_n(1)$, 其中 F_n 为待定的最优控制率. 上述系统的初始值为 $E_n(1) = E(n)$, $d_n(t)$ 和 $L d_n(t)$ 是有界的随机扰动, 且有 $d_n(1) = d(n)$ 和 $|d_n(t)| \leq |d(n)| < \rho$, $t = 1, 2, \dots$, $n = 1, 2, \dots$ (3.19) 对应的名义系统为

$$E_n(t+1) = A_n E_n(t) + B_n \Delta \alpha_n(t), \quad t = 1, 2, \dots \quad (3.20)$$

根据上一小节的分析知, 通过控制名义系统 (3.20) 就能够控制 (3.19). 如果控制策略可以使 (3.20) 在任意的时刻 n 压缩, 那么应用上述控制输入可以使得 (3.19) 中的未知且有界的干扰项被限制在一个紧集中, 且也使得 $e(n)$ 等状态变量收敛到一个较小的紧集中 [60, 61]. 因此, 我们只考虑通过名义系统的镇定来得到学习方法. 应用 MPC 理论, 对于名义系统 (3.20), 在时域 N 上构造如下的最优控制指标:

$$\begin{aligned} V = & \min_{\Delta \alpha_n(1), \Delta \alpha_n(2), \dots, \Delta \alpha_n(N)} \sum_{t=1}^{N=\infty} E_n^T(t) Q E_n(t) + \gamma \Delta \alpha_n^T(t) R \Delta \alpha_n(t), \\ \text{s.t. } & \Delta \alpha_n(t) = F_n C E_n(t), \quad E_n(t+1) = A_n E_n(t) + B_n \Delta \alpha_n(t), \quad t = 1, 2, \dots, \end{aligned} \quad (3.21)$$

其中 F_n 由优化结果所决定, R 是由核 Gram 矩阵定义的正定矩阵. (3.21) 中等式的右边的第一项度量了状态偏差, 第二项度量了控制输入大小, $\gamma > 0$ 为加权参数, 用于衡量状态偏差和控制输入在指标中的相对权重. 为了方便计算, 设 Q 为单位阵 I . 一旦得到了 (3.21) 的解, 我们仅使用第一个控制输入 $\Delta \alpha_n(1)$ 来对模型进行更新, 即参数更新和学习策略由 $\Delta \alpha(n) = \Delta \alpha_n(1)$ 给出. 当下一个时刻 $n+1$ 到来、新的观测值进入学习时, 通过再次求解更新的优化问题来得到 $n+1$ 时刻的控制率 F_{n+1} (A_n 和 B_n 更新为 A_{n+1} 和 B_{n+1}). 随着时间的演进, 通过不断重复上述求解过程, 我们便可以根据所得到的控制率来设计在线学习算法.

定理 3.1 优化问题 (3.21) 的解可以由解如下的关于 P_n 、 S_n 和 F_n 的非线性矩阵方程而得到:

$$\begin{cases} (A_n + B_n F_n C)^T P_n (A_n + B_n F_n C) - P_n + Q + C^T F_n^T R F_n C = 0, \\ (A_n + B_n F_n C) S_n (A_n + B_n F_n C)^T - S_n + I = 0, \\ (R + B_n^T P_n B_n) F_n (C S_n C^T) - B_n^T P_n A_n S_n C^T = 0, \end{cases} \quad (3.22)$$

$F_n = (R + B_n^T P_n B_n)^{-1} B_n^T P_n A_n S_n C^T (C S_n C^T)^{-1}$, 且对 $\alpha(n+1)$, 有

$$\alpha(n+1) = \alpha(n) + \Delta\alpha_n(1) = \alpha(n) + F_n C E_n(1) = \alpha(n) + F_n e(n). \quad (3.23)$$

为求解 (3.22), 我们也有如下的数值算法

(1) 初始化: 令 $j = 0$, 用线性矩阵不等式方法 (linear matrix inequality, LMI) 得到 $F_n(0)$, 使得 $A_n - B_n F_n(0)C$ 渐近收敛.

(2) 第 j 步迭代: 通过解如下矩阵方程得到 $P_n(j)$ 和 $S_n(j)$:

$$\begin{cases} (A_n + B_n F_n(j)C)^T P_n(j) (A_n + B_n F_n(j)C) - P_n(j) + Q + C^T F_n^T(j) R F_n(j) C = 0, \\ (A_n + B_n F_n(j)C) S_n(j) (A_n + B_n F_n(j)C)^T - S_n(j) + I = 0. \end{cases} \quad (3.24)$$

(3) 更新 $F_n(j)$. 令 $F_n(j+1)^* = (R + B_n^T P_n(j) B_n)^{-1} B_n^T P_n(j) A_n S_n(j) C^T (C S_n(j) C^T)^{-1}$, $F_n(j+1) = \lambda F_n(j) + (1-\lambda) F_n(j+1)^*$, 其中 $0 < \lambda < 1$ 且 $A_n + B_n F_n(j+1)C$ 是压缩的.

(4) 如果 $P_n(j)$ 和 $P_n(j+1)$ 足够接近, 即 $\|P_n(j+1) - P_n(j)\|$ 小于一个给定的阈值, 那么停止迭代, $F_n = F_n(j+1)$. 否则重复第 3 步.

上述定理既为误差动力学系统设计了输出反馈控制率的计算方式, 也为学习问题提供了算法基础. 相应的计算可以通过使用 Matlab 中最优化工具箱来实现. 接着考虑基向量 $B_S = \{\phi(u_1), \phi(u_2), \dots\}$ 的选择. 借助于近似线性相关条件, 令 M_0 为 B_S 中向量数目的最大值. 假设在 n 时刻, 我们已经有了参数估计 $w(n) = \sum_{i=1}^M \alpha_i(n) \phi(u_i)$. 对于新进入的样本 $(y(n+1), x(n+1))$, 令

$$\zeta(n) = \min_c \left\| \sum_{i=1}^M c(i) \phi(u_i) - \phi(x(n+1)) \right\|^2, \quad (3.25)$$

$\zeta(n)$ 可以被展开为

$$\zeta(n) = K(x(n+1), x(n+1)) - K_M(x(n+1))^T K_M^{-1} K_M(x(n+1)), \quad (3.26)$$

其中

$$K_M(x(n+1)) = (K(u_1, x(n+1)), K(u_2, x(n+1)), \dots, K(u_M, x(n+1)))^T$$

且 $K_M^{-1} = [K(u_i, u_j)]_{1 \leq i, j \leq M}^{-1}$.

如果 $\zeta(n)$ 较大, 那么由 $\{\phi(u_1), \phi(u_2), \dots, \phi(u_M)\}$ 张成的空间不能很好地线性表示 $\Phi(x(n+1))$, 这意味着该空间对目标函数的近似能力比 $B_S = \{\phi(u_1), \phi(u_2), \dots, \phi(u_M)\} \cup \{\phi(x(n+1))\}$ 张成的空间弱很多, 可能不足以精确地刻画系统的动力学性质. 如果 $\zeta(n)$ 非常小, 那么两个空间近似能力基本相同, $\phi(x(n+1))$ 作为基向量会很冗余. 此外, 对于给定的足够小的常数 ν , 如果 $\zeta(n) < \nu$, 那么在误差 ν 内, $\phi(x(n+1))$ 可以近似表示成基向量 $\{\phi(u_i)\}_{1, \dots, M}$ 的线性组合. 因此, 我们提出如下的基向量在线选择策略. 如果 $\zeta(n) \geq \nu$ 且 $M < M_0$, 则将 $\phi(x(n+1))$ 加入到 B_S 中, B_S 更新为

$\{\phi(u_1), \phi(u_2), \dots, \phi(u_M)\} \cup \{\phi(x(n+1))\}$. 如果 $\zeta(n) < \nu$, 则 B_S 保持不变. 基于这一策略, 选出的基向量 $\phi(u_i)$ 在 RKHS 中将不会出现完全的线性相关. 因此, 可以运用前述的引理, 以保证我们学习方法的合理性. ν 是阈值参数, 它决定了基向量的个数和 H_K^S 在 H_K 中的稀疏性. 一个较小的 ν 可以使学习模型有着更好的逼近性能, 但是也将提高模型计算复杂度, 一个较大的 ν 可以减小计算量, 但也不可能使得模型过于简单, 逼近能力有所下降. 这里还规定, 如果 $\phi(x(n+1))$ 加入到 B_S 中, 我们便将 $w(n) = \sum_{i=1}^M \alpha_i(n)\phi(u_i)$ 重写为 $w(n) = \sum_{i=1}^M \alpha_i(n)\phi(u_i) + 0\phi(u_{M+1})$, 其中 $u_{M+1} = x(n+1)$. 控制系统变为

$$E(n+1) = A(n)E(n) + B(n)\Delta\alpha(n) + Ld(n), \quad (3.27)$$

其中

$$E(n) = (e(n), \overline{\alpha(n)})^T, \quad A(n) = \begin{bmatrix} 1 & \Delta\bar{K}(n) \\ 0_{M+1,1} & I_{M+1,M+1} \end{bmatrix}, \quad B(n) = (\bar{K}(n+1), I_{M+1,M+1})^T,$$

$L = (1, 0_{M+1,1}^T)^T$, $\bar{K}(n+1) = (K(u_1, x(n+1)), \dots, K(u_M, x(n+1)), K(x(n+1), x(n+1)))$. $\alpha(n) = (\alpha_1(n), \alpha_2(n), \dots, \alpha_M(n))$ 重新写成 $\alpha(n) = (\alpha_1(n), \alpha_2(n), \dots, \alpha_M(n), 0)$, 并将根据未来时刻的控制输入进行更新.

综上所述, 我们将新的在线学习算法梳理如下:

- (1) 给定初始值, 基向量 $B_S = \emptyset$, 阈值为 ν 和 M_0 .
- (2) 对于 $n = 1$, 有初始值 $\alpha = 0$, $B_S = \{x(1)\}$ 和 $e(1) = -y(1)$.
- (3) 当 $n > 1$ 时刻, $B_S = \{u_1, u_2, \dots, u_M\}$, 并且由识别参数 $w(n) = \sum_{i=1}^M \alpha_i(n)\phi(u_i)$, $\alpha(n) = (\alpha_1(n), \alpha_2(n), \dots, \alpha_M(n))$. 如果 $\zeta \neq \nu$, 则 B_S 保持不变, $\alpha(n)$ 根据 (3.23) 更新为 $\alpha(n+1)$. 当 $\zeta > \nu$ 且 $M < M_0$ 时, $B_S = \{u_1, u_2, \dots, u_M, u_{M+1}\}$, 这时 $\alpha(n)$ 则通过求解基于 (3.27) 和 (3.23) 的最优化问题更新为 $\alpha(n+1)$.
- (4) 如果数据学习完毕, 则终止程序; 否则令 $n = n+1$ 并重复第 3 步.

3.3 半参数模型的推广

我们的方法也可以用来解决半参数模型的在线学习问题. 对于给定的向量 $x(n)$, 定义一个唯一的向量划分 $x(n) = (x_1(n), x_2(n))$, x_1 和 x_2 分别表示构成回归模型线性和非线性部分的自变量集合. 假设 $x_1(n)$ 是 M_1 维向量, $x_2(n)$ 是 d_1 维向量. 我们有数据流 $(x(1), y(1)), (x(2), y(2)), \dots, (x(n), y(n)), \dots$, 半参数模型可表示为

$$y(n) = x_1(n)\beta + f(x_2(n)) + \epsilon(n), \quad (3.28)$$

其中 β 为线性部分的参数向量, f 为非线性部分. 对 f 和 β 而言, 假设存在理想的目标参数向量 β^* 和 $w^* \in H_K$, 使得

$$y(n) = x_1(n)\beta^* + \phi(x_2(n))w^* + \varepsilon(n), \quad (3.29)$$

对于每一个离散的时间点 n , 有如下学习模型:

$$\hat{y}(n) = x_1(n)\beta(n) + \phi(x_2(n))w(n). \quad (3.30)$$

令 $e(n) = \hat{y}(n) - y(n)$, $\Delta\beta(n) = \beta(n+1) - \beta(n)$, $\Delta x_1(n) = x_1(n+1) - x_1(n)$ 和 $\Delta\phi(x_2(n)) = \phi(x_2(n+1)) - \phi(x_2(n))$, 使用第 3.1 小节中相似的方法, 可以得到半参数模型的误差动力学系统

$$e(n+1) = e(n) + x_1(n+1)\Delta\beta(n) + \phi(x_2(n+1))\Delta w(n) + d^*(n), \quad (3.31)$$

其中

$$d^*(n) = \Delta x_1(n)(\beta(n) - \beta^*) + \Delta\phi(x_2(n))(w(n) - w^*) + \varepsilon(n) - \varepsilon(n+1).$$

注意到 $w(n+1) - w^* = w(n) - w^* + \Delta w(n)$ 和 $\beta(n+1) - \beta^* = \beta(n) - \beta^* + \Delta\beta(n)$, 我们有

$$\begin{aligned} \begin{bmatrix} e(n+1) \\ \beta(n+1) - \beta^* \\ w(n+1) - w^* \end{bmatrix} &= \begin{bmatrix} 1 & \Delta x_1(n) & \Delta\phi(x_2(n)) \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} e(n) \\ \beta(n) - \beta^* \\ w(n) - w^* \end{bmatrix} \\ &+ \begin{bmatrix} x_1(n+1) & \phi(x_2(n+1)) \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \Delta\beta(n) \\ \Delta w(n) \end{bmatrix} + \begin{bmatrix} d(n) \\ 0 \\ 0 \end{bmatrix}, \end{aligned} \quad (3.32)$$

其中 $d(n) = \varepsilon(n) - \varepsilon(n+1)$. 与 (3.9) 中一样, 非线性部分 f 可以由 H_K 中的 M_2 个基向量张成的线性子空间近似得到, 基向量记为 $\phi(u_i)$ ($i = 1, \dots, M_2$), 令 $w(n) = \sum_{i=1}^{M_2} \alpha_i(n)\phi(u_i)$, $w^* = \sum_{i=1}^{M_2} \alpha_i^*\phi(u_i)$, $\overline{\beta(n)} = \beta(n) - \beta^*$, $\overline{\alpha(n)} = (\alpha_1(n) - \alpha_1^*, \dots, \alpha_{M_2}(n) - \alpha_{M_2}^*)^T$, 沿用与 (3.9)–(3.11) 相同的数学表达式, 我们有如下有限维误差系统:

$$\begin{aligned} \begin{bmatrix} e(n+1) \\ \overline{\beta(n+1)} \\ \overline{\alpha(n+1)} \end{bmatrix} &= \begin{bmatrix} 1 & \Delta x_1(n) & \Delta\overline{K(n)} \\ 0 & I_{M_1, M_1} & 0 \\ 0 & 0 & I_{M_2, M_2} \end{bmatrix} \cdot \begin{bmatrix} e(n) \\ \overline{\beta(n)} \\ \overline{\alpha(n)} \end{bmatrix} \\ &+ \begin{bmatrix} x_1(n+1) & \overline{K(n+1)} \\ I_{M_1, M_1} & 0 \\ 0 & I_{M_2, M_2} \end{bmatrix} \cdot \begin{bmatrix} \Delta\beta(n) \\ \Delta\alpha(n) \end{bmatrix} + \begin{bmatrix} d(n) \\ 0 \\ 0 \end{bmatrix}. \end{aligned} \quad (3.33)$$

以上误差动力学系统也可以被写成 (3.12) 的形式. 因此, 本文提出的方法可以被自然地用来解决半参数模型的学习问题.

3.4 关于算法的一些讨论

本文首次将非线性动力学系统的学习问题转化为一组离散线性系统的模型预测最优控制问题, 并且通过重复的求解二次优化指标 (3.21) 来得到控制率和学习算法. 一方面, 尽管对于大多数系统而言, 一个足够大的系数依赖的 N 可以使得通过 (3.21) 得到的闭环系统是稳定的, 但仍然不能完全保证模型预测控制在闭环工作中达到最优控制效果, 在一些特定的情形下闭环系统甚至可能更加不稳定. 如果把 N 取为无穷大, 不仅可以完全解决控制效果和闭环系统的稳定性问题^[39, 62], 而且可以使得对于任意的时刻 n , 指标 (3.21) 有着一致的形式, 这也大大简化了我们求解最优问题的计算. 另一方面, 容易证实, 对于任意给定的 n , (A_n, B_n) 能控, 矩阵 C 行满秩, R 是正定矩阵, (\sqrt{Q}, A) 是可测的. 在这些

条件下, (3.21) 中解的存在性和相应闭环系统的稳定性可以通过文献 [39] 中关于 MPC 的一些经典结果得到保证. 因而, 我们的方法是自适应收敛的, 不需要引入学习步长, 这也是新方法的一个优点.

在基于增量优化的学习方法 [31, 32, 38] 中, 模型参数的更新方向很大程度上取决于性能指标函数. 当数据流中存在噪声影响时, 性能指标函数便不再连续, 这意味着估计参数所在的超平面不再平滑, 如果噪声具有异方差、序列相关或非稳态的性质, 就会使得超平面越发粗糙. 这会导致学习算法所依赖的梯度不一定总能提供正确的参数更新方向, 从而会影响学习效果的鲁棒性和精确性. 本文提出的方法通过控制输入使得误差动力学系统渐近稳定, 因而可以保证在任意性质的噪声扰动下, 估计模型的参数在学习的每一步中都沿着正确的方向更新, 且使得估计的误差收敛到一个较小的有界域, 这大大提高了在线学习的速度、适应性和鲁棒性. 这也是新方法的另一个主要优点.

容易看到, 在 (3.21) 中, γ 越大, $\Delta\alpha$ 越容易被控制和约束, 使得在出现离群值或较大的噪声的情形下, 模型对于动力系统保持相对较慢而稳定的跟踪, 以达到良好的预测效果. 在 γ 较小的情形下, $\Delta\alpha$ 受到的限制相对较小, 会带来较快的学习速度, 但是这也可能会造成过拟合问题, 且模型的泛化能力可能也会降低. 因此, 可以认为权重参数 γ 与传统的核学习方法中的 Tihonov 正则化参数扮演着相似的角色, 它不仅决定了模型复杂度和最小化近似误差之间的权衡, 而且控制了最终解的平滑程度. 因而, 新的方法依然保持了传统核学习方法的一些特点.

另外, 这里虽然仅使用了模型预测控制来得到学习算法, 但可以看到, 基于本文提出的误差动力学系统 (3.12), 诸如 H_∞ 反馈控制、比例 - 积分 - 微分控制 (PID) 和时间最优控制等一些经典的控制方法也可以用于镇定系统 (3.12), 从而可能得到一些新的有效的鲁棒学习算法. 这些有趣的内容我们也会在未来的研究中逐一展开.

4 数值例子及比较

本节用数值例子说明我们提出的基于鲁棒最优控制核学习方法 (robust optimal control based kernel learning method, ROCKL) 的有效性, 并将 ROCKL 与一些经典的核学习算法进行了对比.

4.1 带有非 Gauss 且序列相关随机扰动的非线性动力学系统

研究如下的随机动力学系统 [30, 63]:

$$\begin{cases} y(t) = \frac{y(t-1)y(t-2)y(t-3)(y(t-2)-2)u(t-1) + u(t)}{1 + 5y(t-3)^2} + \varepsilon(t), \\ u(k) = \sin\left(\frac{\pi t}{125}\right), \quad 4 \leq t \leq 1000, \\ y(t) = \frac{y(t-1)y(t-2)y(t-3)(y(t-3)-1)u(t-1) + u(t)}{1 + y(t-2)^2 + y(t-3)^2} + \varepsilon(t), \\ u(k) = \frac{2}{5} \sin\left(\frac{\pi t}{125}\right) + \frac{3}{5} \sin\left(\frac{\pi t}{25}\right), \quad 1001 \leq t \leq 2000, \end{cases} \quad (4.1)$$

$y(1) = 0.5$, $y(2) = 0.55$, $y(3) = 0.6$. 令 \mathcal{B} 为滞后算子. 噪声项 $\varepsilon(t)$ 非 Gauss 且序列相关, $\varepsilon(t) = 0.1(1 - 0.6\mathcal{B})^{-1}(\varepsilon_1(t) + \varepsilon_2(t))$, 其中 $\varepsilon_1(t) \sim N(0, 1)$, $\varepsilon_2(t)$ 在 $[-0.5, 0.5]$ 上服从均匀分布. 上述动力学系统在时刻 $t = 1,000$ 发生突变. 我们生成 2,000 个数据, 如图 1(a) 所示. 在本例中, 学习模型的输入向量由 $u(t)$ 和 $y(t)$ 以及它们的滞后项组成, 即 $x(t) = [y(t-1), y(t-2), y(t-3), u(t)]$. 我们将用 ROCKL 算法得到的结果与用经典算法核最小均方二乘方法 KLMS [31, 32]、核递归最小二乘方法 KRLS (kernel

recursive least square)^[29] 和移动窗框最小二乘支持向量机 (moving window least square support vector machine, MW-LSSVM) 得到的结果进行对比, 以说明新方法的优势. 在每一个算法中, 我们使用具有相同带宽 σ 的 Gauss 核函数 $K(x_i, x_j) = \exp(-\|x_i - x_j\|_2^2/\sigma^2)$, 以得到一致可比的结果.

图 1(b) 给出了当 $\sigma = 0.2$ 、 $\gamma = 0.05$ ((3.21) 中的控制参数) 和 $\nu = 0.5$ (ALD 阈值参数) 时 ROCKL 的输出结果, ROCKL 最后选择了 19 个样本作为基向量 (1, 2, 3, 4, 5, 7, 22, 26, 29, 56, 59, 87, 129, 136, 152, 185, 1055, 1166, 1179). 可以看到尽管系统遭遇了突变, ROCKL 仍然可以很好地学习并预测系统的状态. 图 2(a) 给出了 ROCKL 的预测误差. MW-LSSVM 的预测误差如图 2(b) 所示, 其中正则化参数为 0.01, 支持向量的个数为 $N^* = 10$. KLMS 和 KRLS 的预测误差分别如图 3(a) 和 3(b) 所示, 其中 KLMS 的学习步长 η 为 0.1, 另外, MSE (mean square error, 均方误差) 定义为

$$\text{MSE} = \sqrt{\frac{1}{1970} \sum_{t=31}^{2000} (\hat{y}(t) - y(t))^2},$$

其中 $y(t)$ 的预测值记为 $\hat{y}(t)$, 去掉前 30 个误差数据以排除瞬时响应的影响, 表 1 给出了不同参数下应用不同方法学习所得到的 MSE. 可以看到, 虽然不同的核学习参数的选择会对预测结果造成一定的

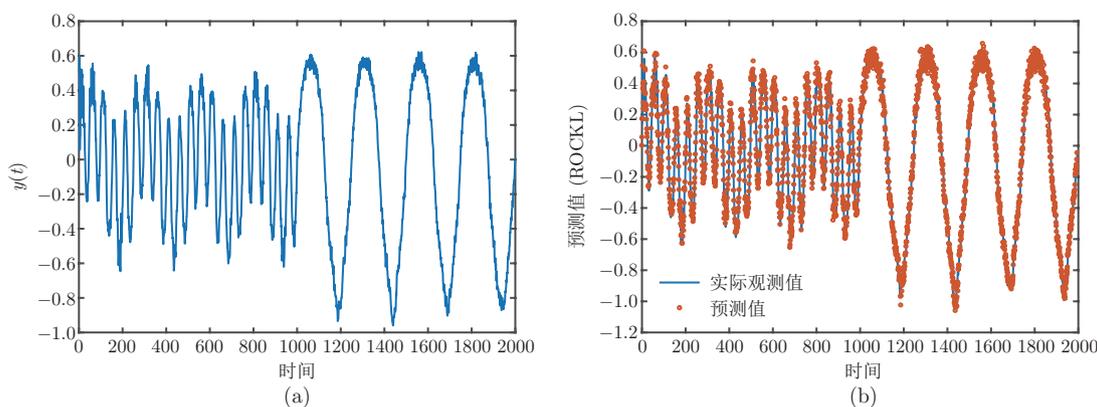


图 1 系统 (4.1) 输出数据与鲁棒最优控制核学习方法预测输出. (a) 系统 (4.1) 模拟输出; (b) 鲁棒最优控制核学习方法预测输出

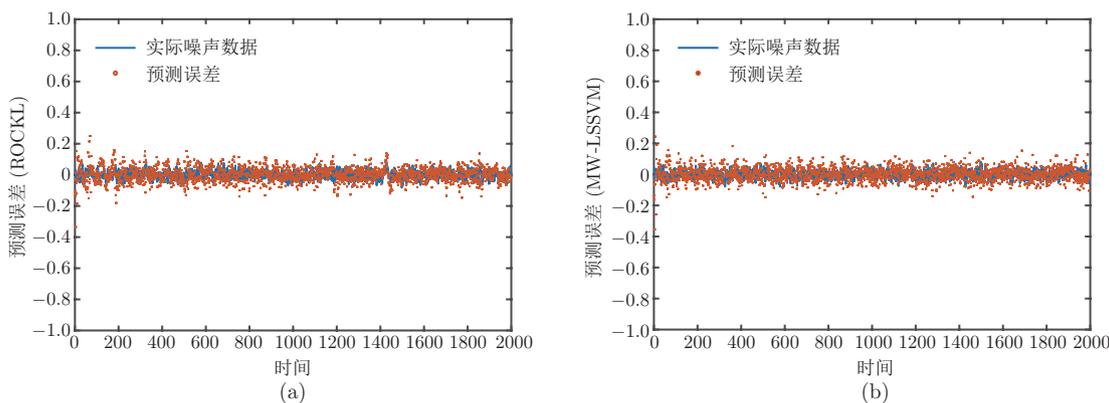


图 2 鲁棒最优控制核学习方法与移动窗框最小二乘支持向量机的预测误差. (a) 鲁棒最优控制核学习方法预测误差; (b) 移动窗框最小二乘支持向量机预测误差

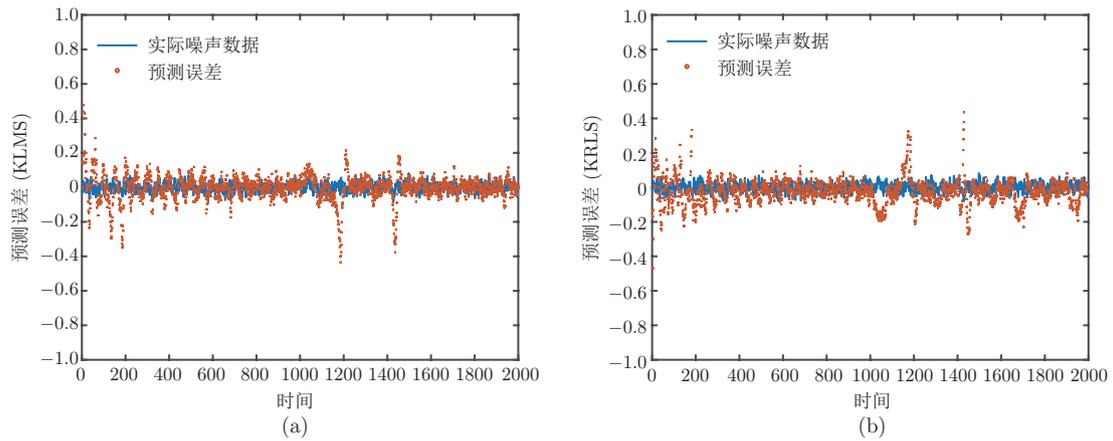


图 3 核最小均方二乘方法与核递归最小二乘方法的预测误差. (a) 核最小均方二乘方法预测误差; (b) 核递归最小二乘方法预测误差

表 1 不同参数的学习结果

方法	σ	γ	ν	η	N^*	MSE
ROCKL	0.1	0.05	0.5			0.0473
MW-LSSVM	0.1	0.05			10	0.0747
KLMS	0.1			0.10		0.0778
KRLS	0.1		0.5			0.0789
ROCKL	0.2	0.10	0.5			0.0419
MW-LSSVM	0.2	0.10			10	0.0869
KLMS	0.2			0.15		0.0631
KRLS	0.2		0.5			0.0764
ROCKL	0.3	0.20	0.5			0.0510
MW-LSSVM	0.3	0.20			10	0.0893
KLMS	0.3	0.20		0.20		0.0814
KRLS	0.3		0.5			0.0721

影响, 但是, 当参数值限制在某一个合理范围内时, 并且当选择相同的核参数 σ 时, ROCKL 相较于 KLMS、KRLS 和 MW-LSSVM, 在预测精度、收敛速度和适应性这三个方面都有着更好的表现. 这也意味着新的方法在不同的再生核空间中都可以更好地对动力学系统进行学习, 并且在复杂的噪声环境下表现出最佳的鲁棒性.

4.2 带有异质性随机扰动的非线性动力学系统

研究如下的非线性系统^[56]:

$$\begin{cases} y(t) = \sin(y(t-1)) + 0.02 \log(|y(t-1)|)\varepsilon(t), & 2 \leq t \leq 200, \\ y(t) = \frac{1.8}{1 + y(t-1)^2} + 0.01|y(t-1)|\varepsilon(t), & 201 \leq t \leq 400, \end{cases} \quad (4.2)$$

其中

$$\varepsilon(t) \sim N(0, 1), \quad y(1) = 0.5.$$

(4.2) 生成 400 个时间序列观测数据, 其如图 4(a) 所示. 系统的动力学性质在时刻 $t = 200$ 有一个突变. 在本例中, 输入变量为 $y(t - 1)$, 我们仍然使用 Gauss 核来对系统建模. 当核参数 $\sigma = 0.3$ 、 $\gamma = 0.1$ 和 $\nu = 0.5$ 时, ROCKL 的预测输出和误差分别如图 4(b) 和 5(a) 所示, 并且有 $MSE = 0.0742$. 图 5(b) 给出了 MW-LSSVM 的预测误差, 其中正则化参数设为 0.1, 支持向量的个数 $N^* = 2$. KLMS 和 KRLS 的误差分别由图 6(a) 和 6(b) 给出, 其中 KLMS 的学习步长 η 设为 0.05. 结果表明, 相比其他的方法, ROCKL 可以更快地收敛, 并且预测误差也更小, 学习效果更好.

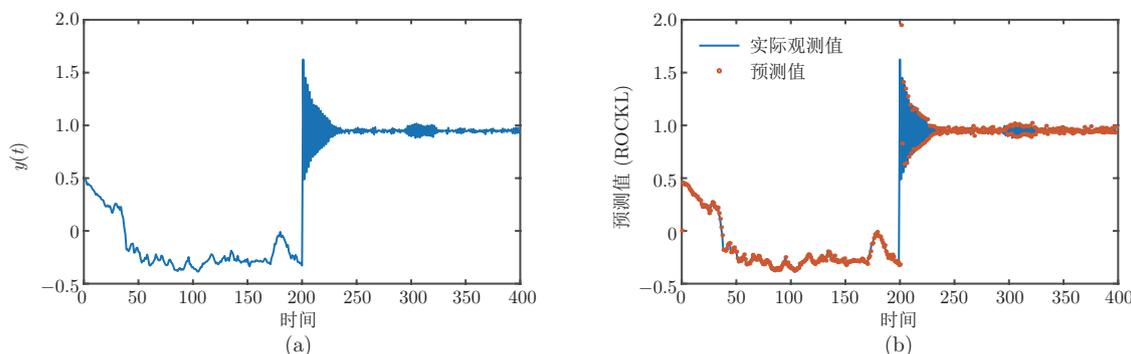


图 4 系统 (4.2) 输出数据与鲁棒最优控制核学习方法预测输出. (a) 系统 (4.2) 的模拟输出; (b) 鲁棒最优控制核学习方法预测输出

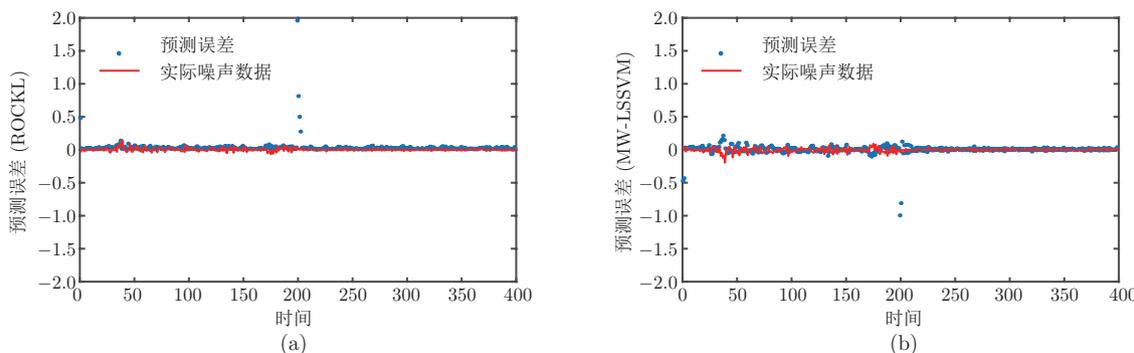


图 5 系统 (4.2) 鲁棒最优控制核学习方法与移动窗框最小二乘支持向量机的预测误差. (a) 鲁棒最优控制核学习方法预测误差; (b) 移动窗框最小二乘支持向量机预测误差

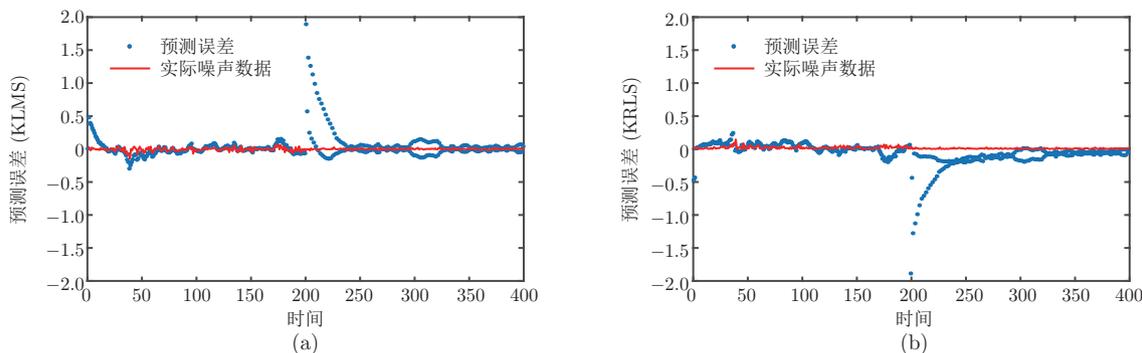


图 6 系统 (4.2) 核最小均方二乘方法与核递归最小二乘方法的预测误差. (a) 核最小均方二乘方法预测误差; (b) 核递归最小二乘方法预测误差

4.3 实际数据分析

本节将所提的方法用于 Box-Jenkins 反应炉二氧化碳浓度时间序列数据的建模中^[64]. 数据可以从 <http://support.sas.com/> 上获得. 我们有数据集 $(u(t), y(t))$, $t = 1, 2, \dots, 296$, $y(t)$ 和 $u(t)$ 分别为反应炉中二氧化碳浓度与天然气气流速度的测量数据, 并分别如图 7(a) 和 7(b) 所示. 令

$$x(t) = (y(t-1), y(t-2), y(t-3), u(t-1), u(t-2), u(t-3))$$

为输入变量, $y(t)$ 为输出变量. 我们使用 Gauss 核来对时间序列进行建模, 当核参数 $\sigma = 20$, $\gamma = 0.1$ 和 $\nu = 0.2$ 时, ROCKL 的预测输出和误差, 以及在线约减最小二乘支持向量回归方法 OR-LSSVR (online reduced least square support vector regression)^[27] 的预测输出和误差分别由图 8 和 9 所示. 显然, 在这个实际数据的例子中, 与 OR-LSSVR 相比, ROCKL 学习效果也更好.

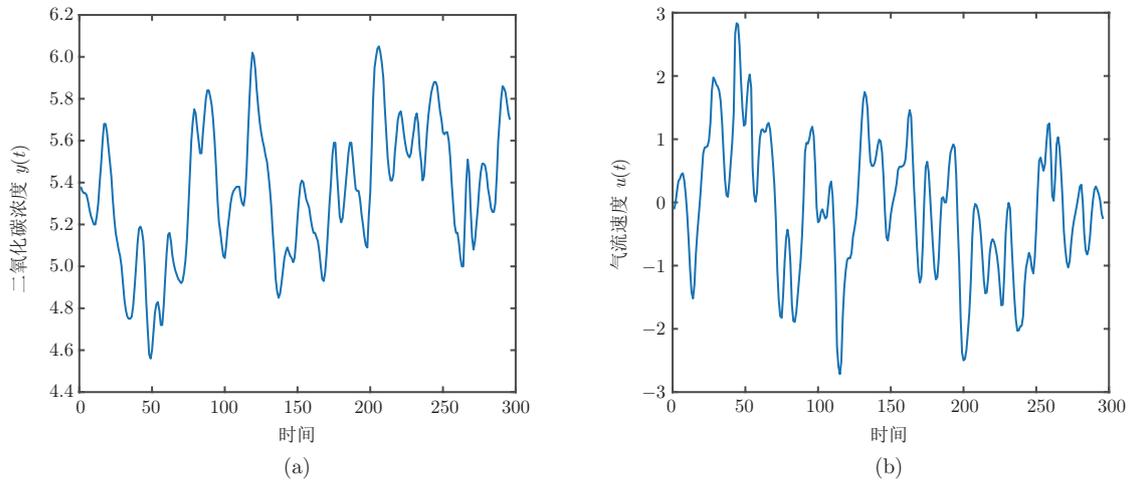


图 7 气体浓度与速度的测量数据. (a) 二氧化碳浓度时间序列数据; (b) 气流速度时间序列数据

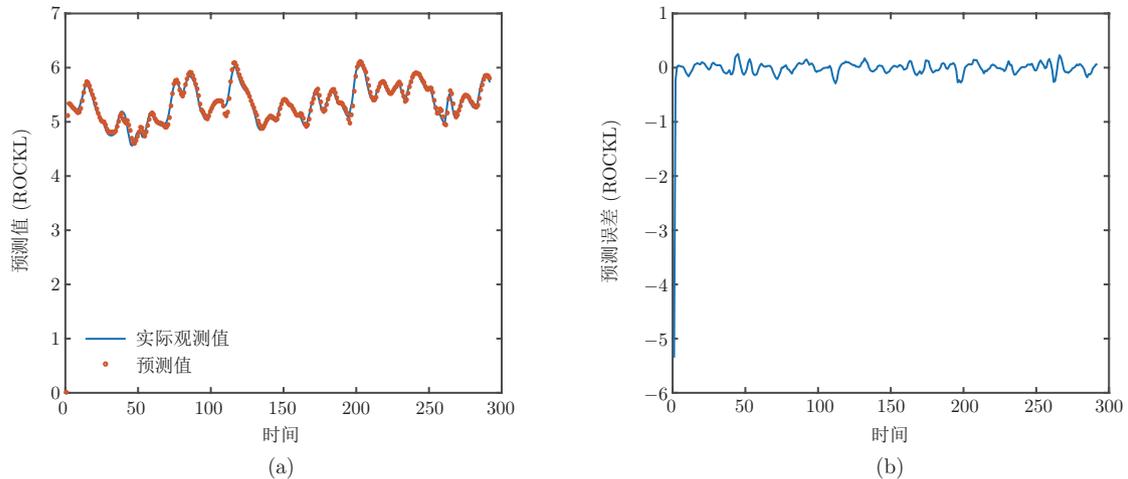


图 8 鲁棒最优控制核学习方法预测输出与预测误差. (a) 鲁棒最优控制核学习方法预测输出; (b) 鲁棒最优控制核学习方法预测误差

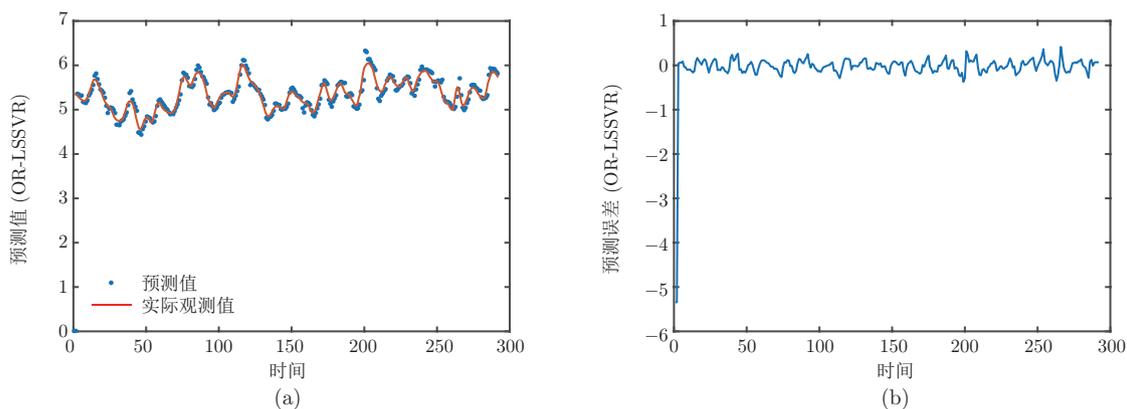


图 9 在线约减最小二乘支持向量回归方法预测输出与预测误差. (a) 在线约减最小二乘支持向量回归方法预测输出; (b) 在线约减最小二乘支持向量回归方法预测误差

4.4 关于算法的进一步讨论

核学习方法中核参数 (带宽参数 σ) 的选择对学习效果有着较大的影响. 一般来讲, 一个小的 σ 对应比较复杂的模型, 在学习非线性性比较强的系统时表现比较好. 一个大的 σ 对应比较平滑的模型, 在学习线性特征占主导的系统时表现较好. 在离线学习中, 最常用且有效的带宽选择策略是交叉验证法, 其基本思想是, 在一定意义下将样本数据集进行分组, 一部分做为训练集, 另一部分做为验证集. 首先用训练集建立估计模型, 再利用验证集来测试估计模型, 以此来设计评价带宽参数的性能指标. 通常来说, 使得验证误差最小的带宽值被认为是最优的. 然而, 在线学习中, 系统的潜在规律可能是时变的, 有效的样本集合难以确定, 而且估计模型需要对系统变化和不断到来的新数据进行及时响应, 因而交叉验证方法并不能被平行地推广过来. 在绝大部分关于在线学习的文献中, 参数依然是被经验性或者启发式的选取, 核参数的适应性计算方法依然是一个公开问题^[30,65]. 最近的一个研究思路是将 σ 视作学习模型中的一个非线性的回归系数, 然后应用梯度方法对它进行更新, 但其依然没有解决学习步长的设置, 以及对噪声的鲁棒性问题^[65]. 在我们未来的研究中, 基于本文提出的思路, 我们将核参数的更新 $\Delta\sigma$ 看作误差动力学系统的非线性控制项, 应用非线性系统鲁棒最优控制技术去尝试解决这个问题. 另外, 我们也对新方法的计算复杂度做一个简要的分析. ROCKL 的计算量主要来自求解与优化问题 (3.21) 相应的矩阵方程 (3.22). 如果通过本文所示的迭代法算法解 Riccati 方程组 (3.22), 在每个时刻 n , 其计算复杂度是 $L_0 o((M+1)^3)$ ^[39,40], 其中 M 为基向量集合 B_s 中元素的个数, L_0 是一个与阈值有关但是与 M 独立的常数. KRLS、MW-LSSVM 和 OR-LSSVR^[64] 的计算复杂度不会超过 $o((M)^2)$. 虽然与这些常用的核学习算法相比, 新方法的复杂度有所增加, 但对于给定的动力学系统是存在上界的, 因而, ROCKL 是适合在线学习的.

5 结论

本文提出了一个基于鲁棒最优控制的在线核学习方法, 并将其用于解决非线性随机动力学系统在在线建模问题. 与现有的方法相比, 新的方法收敛更快, 可以更好地实现对动力学系统的自适应学习, 而且对噪声数据更加鲁棒, 同时并不需要引入任何复杂的样本选择策略. 此外, 本文首次从最优控制的角度来研究动力学系统的建模问题, 这也为学习理论的研究和算法的开发提供了一些新的思路.

6 定理的证明

6.1 引理 3.1 的证明

令 \bar{B} 为紧集 $B \subset \mathbb{R}^d$ 的凸包, Φ 为特征映射. 易知 \bar{B} 是有界的. 如果集合 $\{\Phi(s) : s \in B\}$ 也是一个紧集, 我们需要证明对 $\forall \epsilon_\star > 0$, 存在有限个子覆盖, 使得这些子覆盖的并集可以覆盖 $\{\Phi(s) : s \in B\}$. 因为 Mercer 核函数 $K(s, t)$ 连续且足够光滑, $K(s, t)$ 与它的偏导 $\frac{\partial K(s, t)}{\partial t}$ 和 $\frac{\partial K(s, t)}{\partial s}$ 在 B 和 \bar{B} 上一致有界, 这意味着存在一个常数 C_B 使得

$$K(s, t) < C_B, \quad \frac{\partial K(s, t)}{\partial s} < C_B, \quad \frac{\partial K(s, t)}{\partial t} < C_B, \quad \forall s, t \in \bar{B}. \quad (6.1)$$

对于任意的 $x_0, y_0 \in B$, 由中值定理, 有

$$\|\Phi(x_0) - \Phi(y_0)\|^2 = \frac{\partial K(x_0, y_0^\star)}{\partial t}(x_0 - y_0) - \frac{\partial K(x_0^\star, y_0)}{\partial s}(x_0 - y_0) \leq 2C_B \|x_0 - y_0\|, \quad (6.2)$$

这说明 $\|\Phi(x_0) - \Phi(y_0)\| \leq \sqrt{2C_B \|x_0 - y_0\|}$, 其中 x_0^\star 和 y_0^\star 属于 \bar{B} . Φ 可以被视为一个等度连续映射. 因为 B 是有界紧集, 那么对于 $\forall \epsilon_\star > 0$ 和正的常数 $\epsilon_\star^2/(2C_B)$, \mathbb{R}^d 中存在有限个子覆盖球 $B_i = \{x : \|x - v_i\| < \epsilon_\star^2/(2C_B)\}$, $i = 1, 2, \dots, M_B$ 使得 $B \subset \bigcup_{1 \leq i \leq M_B} B_i$, 其中 v_i 是这些子覆盖的中心.

此外, 可以看到 $\{\Phi(s) : s \in B\} \subset \bigcup_{1 \leq i \leq M_B} \{\Phi(s) : s \in B_i\}$, 并且对于每一个 $\{\Phi(s) : s \in B_i\}$, $\forall \Phi(x) \in \{\Phi(s) : s \in B_i\}$, $\Phi(v_i) \in \{\Phi(s) : s \in B_i\}$, 我们有

$$\|\Phi(x) - \Phi(v_i)\| \leq \sqrt{2C_B \|x - v_i\|} \leq \sqrt{\frac{2C_B \epsilon_\star^2}{2C_B}} = \epsilon_\star. \quad (6.3)$$

因此, 可以得到结论, 对于给定的 $\forall \epsilon_\star > 0$, 存在一个有限的 ϵ_\star 网 $\bigcup_{1 \leq i \leq M_B} \{\Phi(s) : s \in B_i\}$, 使得 $\{\Phi(s) : s \in B\}$. 故 $\{\Phi(s) : s \in B\}$ 是一个紧集, Φ 为一个紧算子. 此外, 令 $M = M_B$, $u_i = v_i$, 该定理的剩余部分也可以自然得到保证. 证毕.

6.2 引理 3.2 的证明

将核 Gram 矩阵 $[K(u_i, u_j)]_{1 \leq i, j \leq M}$ 记为 Ω , RKHS 中的 $(\phi(u_1), \phi(u_2), \dots, \phi(u_M))$ 记为 Φ_M . 对任意非零向量 $\bar{b} = (b_1, b_2, \dots, b_M)$, 有

$$\bar{b}\Omega\bar{b}^T = \bar{b}\Phi_M^T\Phi_M\bar{b}^T = \left\langle \sum_{i=1}^M b_i\phi(u_i), \sum_{i=1}^M b_i\phi(u_i) \right\rangle = \left\| \sum_{i=1}^M b_i\phi(u_i) \right\|^2. \quad (6.4)$$

因为不存在非零向量 (b_1, b_2, \dots, b_M) 使得 $\sum_{j=1}^M b_j\phi(u_j) = 0$, 那么 $\bar{b}\Omega\bar{b}^T > 0$, 这说明 Ω 为正定矩阵. 此外, 如果存在不同的 c_1 和 c_2 , 使得

$$\sum_{i=1}^M c_1(i)\phi(u_i) = \sum_{i=1}^M c_2(i)\phi(u_i),$$

令 $c^\star = c_1 - c_2$, 且 c^\star 为非零向量, 则有 $\sum_{i=1}^M c_i^\star\phi(u_i) = 0$. 这与引理中的假设矛盾. 证毕.

6.3 定理 3.1 的证明

假设在时刻 n , 由求解 (3.21) 得到的闭环系统为

$$E_n(t+1) = (A_n + B_n F_n C) E_n(t). \quad (6.5)$$

性能指标可以表示成包含反馈控制率矩阵 F_n 的形式

$$V = \sum_{t=1}^{N=\infty} E_n^T(t) (Q + \gamma C^T F_n R F_n C) E_n(t). \quad (6.6)$$

通过 Lyapunov 方法, 可以将动力学系统的镇定问题转为一个等价形式. 设存在一个对称正定矩阵 P_n , 使得

$$E_n^T(t+1) P_n E_n(t+1) - E_n^T(t) P_n E_n(t) = -E_n^T(t) (Q + \gamma C^T F_n R F_n C) E_n(t), \quad (6.7)$$

那么, (6.6) 可以写成

$$\begin{aligned} V &= \sum_{t=1}^{N=\infty} E_n^T(t) P_n E_n(t) - E_n^T(t+1) P_n E_n(t+1) \\ &= E_n^T(1) P_n E_n(1) - \lim_{t \rightarrow \infty} E_n^T(t) P_n E_n(t). \end{aligned} \quad (6.8)$$

假设闭环系统渐近稳定, 那么随着 t 增加, $E_n(t)$ 渐近收敛到 0, 因此, $V = E_n^T(1) P_n E_n(1)$. 在另一方面, 通过 (6.5) 和 (6.7), 可得

$$E_n^T(t+1) P_n E_n(t+1) - E_n^T(t) P_n E_n(t) = -E_n^T(t) (Q + \gamma C^T F_n R F_n C) E_n(t), \quad (6.9)$$

由于必须满足所有的初始值, 这意味着

$$(A_n + B_n F_n C)^T P_n (A_n + B_n F_n C) - P_n + Q + \gamma C^T F_n R F_n C = 0. \quad (6.10)$$

上式可以被用于计算反馈矩阵 F_n . 性能指标也可以由以下述的形式给出:

$$V = \frac{1}{2} \text{tr}(P_n \mathcal{E}_n), \quad (6.11)$$

其中系统矩阵 $\mathcal{E}_n \equiv E_n(1) E_n^T(1)$. 显然, 计算 $\Delta \alpha(n)$ 的输出反馈控制问题等价于选择 F_n 用以最小化 (6.6) 的优化问题, 其中 (6.11) 受到含 P_n 的辅助矩阵 (6.10) 的约束. 使用 Lagrange 乘数法求解这个问题, 将约束增加到性能指标上, 定义如下 Hamilton 函数:

$$\mathcal{L}_n = \text{tr}(P_n \mathcal{E}_n) + \text{tr}[(A_n + B_n F_n C)^T P_n (A_n + B_n F_n C) - P_n + Q + \gamma C^T F_n R F_n C] S_n, \quad (6.12)$$

其中 S_n 为待计算的 Lagrange 乘数系统矩阵. 令 \mathcal{L}_n 对所有自变量 P_n 、 S_n 和 F_n 的偏导均为零, 含有 F_n 的优化问题的解可以由下式给出:

$$\begin{aligned} \frac{\partial \mathcal{L}_n}{\partial S_n} &= (A_n + B_n F_n C)^T P_n (A_n + B_n F_n C) - P_n + Q + \gamma C^T F_n R F_n C = 0, \\ \frac{\partial \mathcal{L}_n}{\partial P_n} &= (A_n + B_n F_n C) S_n (A_n + B_n F_n C)^T - S_n + \mathcal{E}_n = 0, \end{aligned}$$

$$\frac{\partial \mathcal{L}_n}{\partial F_n} = (R + B_n^T P_n B_n) F_n (C S_n C^T) - B_n^T P_n A_n S_n C^T = 0.$$

可以求解得到 $F_n = (R + B_n^T P_n B_n)^{-1} B_n^T P_n A_n S_n C^T (C S_n C^T)^{-1}$. \mathcal{E}_n 的存在使得优化问题的解依赖系统的初始状态. 这种依赖性输出反馈控制的典型特征. 文献 [39, 58] 建议, 在每个时刻 n 中假定 \mathcal{E}_n 均匀地分布在单位球面上, 从而减轻优化问题对初始值的依赖性, 即设 $\mathcal{E}_n = I$. 这是一个合理的假设, 因为我们试图控制并驱动误差动力学系统, 它可以由任意的非零初始状态逐步收敛到原点.

此外, 容易证实, 矩阵 C 行满秩, 控制权重矩阵 R 正定, (I, A_n) 可检测. 利用文献 [39, 58, 59] 中的结论, 定理提出的算法可以最小化性能指标, 并得到一个稳定闭环系统的输出反馈控制矩阵 F_n . 证毕.

致谢 作者感谢审稿人提出的宝贵建议.

参考文献

- 1 Prato G D, Zabczyk J. Stochastic Equations in Infinite Dimensions. Cambridge: Cambridge University Press, 1992
- 2 Chow P. Stochastic Partial Differential Equations. Boca Raton: Chapman and Hall, 2007
- 3 Peszat S, Zabczyk J. Stochastic Partial Differential Equations with Levy Noise: An Evolution Equation Approach. Cambridge: Cambridge University Press, 2007
- 4 Ma J, Mark W. Recent Advances in Estimating Nonlinear Models: With Applications in Economics and Finance. New York: Springer, 2014
- 5 Li J, Hao W, Zhang X. Learning kernel subspace for face recognition. Neurocomputing, 2015, 151: 1187–1197
- 6 Hagenblad A, Ljung L, Wills A. Maximum likelihood identification of Wiener models. Automatica, 2008, 44: 2697–2705
- 7 Guo J, Wang L Y, Yin G, et al. Identification of Wiener systems with quantized inputs and binary-valued output observations. Automatica, 2017, 78: 280–286
- 8 Ning H, Jing X, Cheng L. Identification of non-linear stochastic spatiotemporal dynamical systems. IET Control Theory Appl, 2013, 17: 2069–2083
- 9 Ning H W, Qing G Y, Jing X J. Identification of nonlinear spatiotemporal dynamical systems with nonuniform observations using reproducing-kernel-based integral least square regulation. IEEE Trans Neural Netw Learn Syst, 2016, 27: 2399–2412
- 10 Scholkopf B, Smola A J. Learning with Kernels: Support Vector Machines, Regularization, Optimization and Beyond. Cambridge: MIT Press, 2002
- 11 Milanese M, Novara C. Set membership identification of nonlinear systems. Automatica, 2004, 40: 957–975
- 12 Jing X, Cheng L. An optimal PID control algorithm for training feedforward neural networks. IEEE Trans Ind Electron, 2013, 60: 2273–2283
- 13 Wei H L, Billings S A, Zhao Y F, et al. An adaptive wavelet neural network for spatio-temporal system identification. Neural Netw, 2010, 23: 1286–1299
- 14 Chen S, Billings S A, Luo W. Orthogonal least squares methods and their application to non-linear system identification. Internat J Control, 1989, 50: 1873–1896
- 15 Chen S, Xia H, Harris C J. Particle swarm optimization aided orthogonal forward regression for unified data modeling. IEEE Trans Evol Comput, 2010, 14: 477–499
- 16 Berlinet A, Christine T A. Reproducing Kernel Hilbert Spaces in Probability and Statistics. Boston: Kluwer Academic Publishers, 2004
- 17 Liu W, Principe J C, Haykin S. Kernel Adaptive Filtering. New York: Wiley, 2010
- 18 Chen X G, Gao Y, Wang R L. Online selective kernel-based temporal difference learning. IEEE Trans Neural Netw Learn Syst, 2013, 24: 1944–1956
- 19 Vapnik V. The Nature of Statistical Learning Theory. New York: Springer-Verlag, 1995
- 20 Suykens J A K, Gestel V T, Brabanter J D, et al. Least Squares Support Vector Machines. Singapore: World Scientific, 2002
- 21 Xu Y L, Chen D R, Li H X, et al. Least square regularized regression in sum space. IEEE Trans Neural Netw Learn Syst, 2013, 24: 635–646
- 22 Ning H, Jing X. Identification of partially known non-linear stochastic spatio-temporal dynamical systems by using a novel partially linear Kernel method. IET Control Theory Appl, 2015, 9: 21–33
- 23 Mehmet G, Ethem A. Multiple kernel learning algorithms. J Mach Learn Res, 2011, 12: 2211–2268
- 24 Lee Y J, Huang S Y. Reduced support vector machines: A statistical theory. IEEE Trans Neural Netw, 2007, 18: 1–13

- 25 Richard C, Bermudez J C M, Honeine P. Online prediction of time series data with kernels. *IEEE Trans Signal Process*, 2009, 57: 1058–1067
- 26 Ning H W, Jing X J, Cheng L. Online identification of nonlinear spatiotemporal systems using kernel learning approach. *IEEE Trans Neural Netw*, 2011, 22: 1381–1394
- 27 Zhao Y P, Sun J G, Du Z H, et al. Online independent reduced least squares support vector regression. *Inform Sci*, 2012, 201: 37–52
- 28 de Kruif B J, de Vries T J A. Pruning error minimization in least squares support vector machines. *IEEE Trans Neural Netw*, 2003, 14: 696–702
- 29 Engel Y, Mannor S, Meir R. The kernel recursive least-squares algorithm. *IEEE Trans Signal Process*, 2004, 52: 2275–2285
- 30 Liu Y, Wang H, Yu J, et al. Selective recursive kernel learning for online identification of nonlinear systems with NARX form. *J Process Control*, 2010, 20: 181–194
- 31 Liu W, Pokharel P P, Principe J C. The kernel least-mean-square algorithm. *IEEE Trans Signal Process*, 2008, 56: 543–554
- 32 Kivinen J, Smola A J, Williamson R C. Online learning with kernels. *IEEE Trans Signal Process*, 2004, 52: 2165–2176
- 33 Wu Q, Ying Y M, Zhou D X. Learning rates of least-square regularized regression. *Found Comput Math*, 2006, 6: 171–192
- 34 Cucker F, Smale S. On the mathematical foundations of learning. *Bull Amer Math Soc*, 2002, 39: 1–49
- 35 De Brabanter K, De Brabanter J, Suykens J A K, et al. Kernel regression in the presence of correlated errors. *J Mach Learn Res*, 2011, 12: 1955–1976
- 36 Espinoza M, Suykens J A K, De Moor B D. LS-SVM regression with autocorrelated errors. *IFAC Proc Volumes*, 2006, 39: 582–587
- 37 Xie T, Yu H, Wilamowski B. Fast and efficient second-order method for training radial basis function networks. *IEEE Trans Neural Netw Learn Syst*, 2012, 23: 609–619
- 38 Chen B, Liang J, Zheng N, et al. Kernel least mean square with adaptive kernel size. *Neurocomputing*, 2016, 191: 95–106
- 39 Lewis F L, Vrabie D, Syrmos V L. *Optimal Control*. New Jersey: John Wiley and Sons, 2012
- 40 Grune L, Pannek J. *Nonlinear Model Predictive Control: Theory and Algorithms*. Berlin: Springer, 2011
- 41 Lin J, Rosasco L, Zhou D X. Iterative regularization for learning with convex loss functions. *J Mach Learn Res*, 2016, 17: 1–38
- 42 Owens D H. *Iterative Learning Control: An Optimization Paradigm*. Berlin: Springer, 2016
- 43 Aliyu M D S. *Nonlinear \mathcal{H}_∞ -Control, Hamiltonian Systems and Hamilton-Jacobi Equations*. London: CRC Press, 2011
- 44 Graf J. *PID Control Fundamentals*. Scotts Valley: CreateSpace Independent Publishing Platform, 2016
- 45 Duan Z, Wang J, Yang Y, et al. Frequency-domain and time-domain methods for feedback nonlinear systems and applications to chaos control. *Chaos Solitons Fractals*, 2009, 40: 848–861
- 46 Yu H. Approximation of time optimal controls for heat equations with perturbations in the system potential. *SIAM J Control Optim*, 2014, 52: 1663–1692
- 47 Berlinet A, Christine T A. *Reproducing Kernel Hilbert Spaces in Probability and Statistics*. Boston: Kluwer Academic Publishers, 2004
- 48 Liu X Q, Kruger U, Littler T, et al. Moving window kernel PCA for adaptive monitoring of nonlinear processes. *Chemom Intell Lab Syst*, 2009, 96: 132–143
- 49 Liu W F, Park I, Wang Y W, et al. Extended kernel recursive least squares algorithm. *IEEE Trans Signal Process*, 2009, 57: 3801–3814
- 50 Tang H S, Xue S T, Chen R, et al. Online weighted LS-SVM for hysteretic structural system identification. *Eng Struct*, 2006, 28: 1728–1735
- 51 Kushner H J, Yin G G. *Stochastic Approximation and Recursive Algorithms and Applications (Stochastic Modelling and Applied Probability)*. Berlin: Springer, 2003
- 52 Smale S, Yao Y. Online learning algorithms. *Found Comput Math*, 2006, 6: 145–170
- 53 Dong X M, Chen D R. Learning rates of gradient descent algorithm for classification. *J Comput Appl Math*, 2009, 224: 182–192
- 54 Haykin S. *Neural Networks and Learning Machines*. London: Pearson, 2008
- 55 Boyd S, Vandenberghe L. *Convex Optimization*. Cambridge: Cambridge University Press, 2004
- 56 De Brabanter K, De Brabanter J, Suykens J A K, et al. Approximate confidence and prediction intervals for least squares support vector regression. *IEEE Trans Neural Netw*, 2011, 22: 110–120

- 57 Wei H L, Billings S A. Model structure selection using an integrated forward orthogonal search algorithm assisted by squared correlation and mutual information. *Int J Model Identif Control*, 2008, 3: 341–356
- 58 Rawlings J B, Mayne D Q. *Model Predictive Control: Theory and Design*. London: Nob Hill Publishing, 2009
- 59 Wang L. *Model Predictive Control System Design and Implementation Using Matlab*. Berlin: Springer, 2009
- 60 Chisci L, Rossiter J A, Zappa G. Systems with persistent disturbances: Predictive control with restricted constraints. *Automatica*, 2001, 37: 1019–1028
- 61 Li H, Shi Y. Robust distributed model predictive control of constrained continuous-time nonlinear systems: A robustness constraint approach. *IEEE Trans Automat Control*, 2014, 59: 1673–1678
- 62 Anderson B D O, Moore J B. *Optimal Control: Linear Quadratic Methods* (Dover Books on Engineering). Netherland: Dover Publications, 2007
- 63 Wang L X. *Adaptive Fuzzy Systems and Control: Design and Stability Analysis*. Upper Saddle River: Prentice Hall, 1994
- 64 Box G E P, Jenkins G M, Reinsel G C. *Time Series Analysis: Forecasting and Control*, 4th ed. New York: John Wiley and Sons, 2013
- 65 Jing L, Steven C H H, Wang J, et al. Large scale online kernel learning. *J Mach Learn Res*, 2016, 17: 1–43

An adaptive online machine learning method based on a robust optimal control approach

Hanwen Ning & Zhanfeng Li

Abstract Reproducing kernel Hilbert space (RKHS) based models are promising ones for image processing, function approximation, pattern recognition, data mining problems and also have shown their effectiveness in the system identification of nonlinear stochastic dynamical systems. In this paper, a novel control approach to the online learning (regression) problems of RKHS based models is studied in order to develop efficient algorithms with real time and adaptive parameter updates. To this aim, the learning problem for stochastic dynamical systems is reasonably translated into an output feedback control problem for discrete time varying linear dynamical systems with bounded random disturbances by some established new results for RKHS, and an adaptive robust control algorithm is therefore developed for the learning problem using the robust optimal model predictive control techniques. Compared with the existing online kernel learning methods, the proposed one can realize real time model parameter update without introducing any data window principle, pruning technique, adjusting of learning steps and any assumptions on random noise to achieve accurate online modeling performance for stochastic dynamics with abrupt changes, and meanwhile guarantee the fast and robust convergence. Moreover, this study could be the first attempt to use a kernel method to tackle the online learning problems from the perspective of robust optimal control theory. And under the proposed learning framework, existing well established control techniques can be potentially utilized to develop new robust learning methods, resultantly some novel insight for kernel learning theory is provided as well. Theoretical analysis, numeral examples and comparisons are also given to demonstrate our results.

Keywords statistical learning, online learning, reproducing kernel Hilbert space, stochastic dynamical systems, robust optimal control, model predictive control

MSC(2010) 62J02, 68T05, 93A30, 93E35

doi: 10.1360/SCM-2017-0308