第 31 卷 第 6 期 自 动 化 学 报 Vol. 31, No. 6
2005 年 11 月 ACTA AUTOMATICA SINICA November, 2005

学习控制的现状与展望[1)]

许建新[1] 侯忠生[2]
[1](新加坡国立大学电气与计算机系)
[2](北京交通大学电子信息工程学院先进控制系统研究所 北京 100044)
(E-mail: elexujx@nus.edu.sg, houzhongsheng@china.com)

摘 要 在此我们综述学习控制理论的发展历程. 首先给出基于可重复性的学习控制的基本概念与原理. 然后简述学习控制的两大主流 — 迭代学习控制和反复控制. 接着对当今迭代学习控制理论前沿的几大热点加以考察, 并简要介绍学习控制理论和 Lyapunov 方法的结合. 最后讨论学习控制理论的新发展方向.

关键词 学习控制, 迭代学习控制, 反复控制, 综述
中图分类号 TP391.4

On Learning Control: The State of the Art and Perspective

XU Jian-Xin[1] HOU Zhong-Sheng[2]
[1](*Department of Electrical and Computer Engineering, National University of Singapore,*
4 *Engineering Drive* 3, 117576 *Singapore*)
[2](*Advanced Control Systems Laboratory of the School of Electronics and Information Engineering,*
Beijing Jiaotong University, Beijing 100044)
(E-mail: elexujx@nus.edu.sg, houzhongsheng@china.com)

Abstract In this article we review the recent advances in learning control theory. First insight is provided into the concept and principle associated with learning control under a repeatable control environment. Then a briefing is given to the two main streams of learning control-iterative learning control and repetitive control. Subsequently we survey several hot topics in the field of learning control. New trends in learning control using Lyapunov methods are also sketched, followed by remarks on future directions.

Key words Learning control, iterative learning control, repetitive control, survey

1 学习控制的概念

当今控制界理论研究的主流可概括为鲁棒控制 (Robust control) 与智能控制 (Intelligent control). 鲁棒控制本质上是反馈的延伸. 智能控制包罗万象, 至今尚未有明确统一的定义. 笔者认为, 智能控制从狭义上看是前馈的延伸, 从广义上看则应当是具有学习或自我调节功能的控制. 让我们考察一实例: 起重机的操作员. 经过系统培训与自我练习, 即通

1) 国家自然科学基金 (60474038) 资助
Supported by National Natural Science Foundation of P. R. China (60474038)
收稿日期 2004-9-1 收修改稿日期 2005-7-1
Received September 1, 2004; in revised form July 1, 2005

过指导性学习与非指导性学习 (Supervised learning and unsupervised learning), 操作员从一窍不通逐渐成为熟练的技师. 显然学习在这过程中起了至关重要的作用. 遗憾的是, 理论界尚未能建立起完整的关于人的学习机理和学习过程的模型与模式. 专家系统或模糊控制系统跳过学习过程, 直接将人的经验总结为适合机器控制的规则. 因而专家系统或模糊控制系统从狭义上是智能控制, 从广义上看则难以归类为智能控制, 因为其实际学习过程是由人实现的.

学习本是一跨学科的概念, 涉及心理学, 生理学, 生物学, 脑科学, 以及一切和生命有关的学科. 在数理与工程领域我们将学习的概念推广到机器学习 (Machine learning). 当我们着眼于控制理论与控制工程, 有必要给出更具有针对性的学习定义. 控制中的学习大致可分为三个层次: 单一目标的精确学习, 基于模式的多目标学习 (统计学习), 以及量化的生物学习 (Bio-learning). 这三层次的学习与控制相结合构成了学习控制全体. 在此我们只考察单一目标的精确学习控制.

当我们将控制任务分解落实到最基本的执行器, 如阀门、开关、马达、或电流、电压源等, 每一个执行器的任务显然是明确唯一的, 如实现一阶跃或正弦响应. 控制的首要任务就是精确地提供所需的控制信号给执行器, 从而实现单一目标的追踪. 当系统的精确模型已知且逆系统可解, 控制信号可直接算出, 无需学习. 当系统模型并非精确但低增益反馈满足控制要求, 也无需学习. 除此之外, 学习控制就大有用武之地. 注意我们的任务是找出所需的控制信号, 而非学出系统的精确模型. 在这阶段我们有意避免建模, 因为建模过程往往比控制任务本身还要来得复杂. 既然控制目标明确唯一, 我们可构造如下的学习控制模式. 通过执行一次控制任务, 获得了相应控制信号 $u_i(t)$ 与追踪误差信号 $e_i(t)$. 当再次执行同一控制任务时, 在反馈的基础上, 再加上前次的控制信号, 即

$$u_i(t) = u_{i-1}(t) + qe_i(t) \tag{1}$$

q 是一反馈增益, 也是学习增益. 下图显示了学习控制系统的结构.

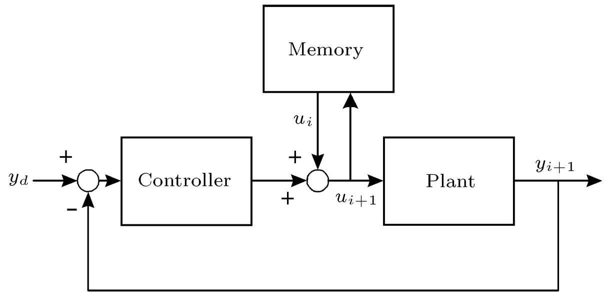

图 1 现行学习控制的结构

Fig. 1 Configuration of current cycle learning

让我们进一步考察这一极其简单的控制算法, 探讨将其称之为学习控制的理由. 先看纯反馈的情况 $u_i(t) = qe_i(t)$. 在控制目标与系统不变的情况下, 重复执行控制任务导致相同的系统响应. 纯反馈无法改善系统响应, 因为控制器未能从过去的执行过程里 “学” 到任何有用信息. 学习控制算法 (1) 则建立了本次与前次控制之间的联系. 可以直观地设想: 当前次的控制信号 u_{i-1} 已充分逼近所需之值, 则 $u_{i-1}(t)$ 作为本次控制信号 $u_i(t)$ 的一组成部分已能胜任, 从而本次误差信号 $e_i(t)$ 充分小, 甚至无需反馈信号系统也能正常工作.

可以看到，$u_{i-1}(t)$ 在 $u_i(t)$ 中扮演了前馈角色．若 u_{i-1} 无法独当一面，本次误差信号 $e_i(t)$ 变大，反馈项 $qe_i(t)$ 立刻投入战斗，支援前馈．$qe_i(t)$ 同时也扮演了误差修正的角色，修正后的控制信号即为 $u_i(t)=u_{i-1}(t)+qe_i(t)$. 从学习的角度看，过去的控制经验 $u_{i-1}(t)$ 被有效地利用以弥补控制先验知识的不足．这和我们人的简单学习过程是相似的．我们总是通过不断重复与修正某一动作，从而掌握正确的做法．又通过成百上千次的强化学习，有意识的动作会转化为下意识或本能的动作．从控制观点看，反复学习导致了内模的建立．控制过程由反馈主导通过学习转化为前馈唱主角．

让我们更进一步延拓学习控制律．先考察学习控制律 (1) 的局限性．首先，学习控制律 (1) 没有充分利用前次控制获得的信息，如前次的追踪误差 $e_{i-1}(t)$. 其次，q 既是反馈增益，也是学习增益．反馈的目的是要保证系统沿时间轴 t 稳定或收敛，学习的目的是要保证系统沿学习轴 i 收敛．可是我们只有一个可调参数 q, 难以兼顾反馈稳定与学习收敛．于是又有了更进一步的大胆设想：干脆去掉去反馈，让前馈式学习控制一统天下 (见 Fig.2)

$$u_i(t)=u_{i-1}(t)+qe_{i-1}(t) \tag{2}$$

只要受控对象的初始状态和动力系统不变，控制目标不变，系统沿学习轴 i 收敛一定能保证沿时间轴 t 的收敛．

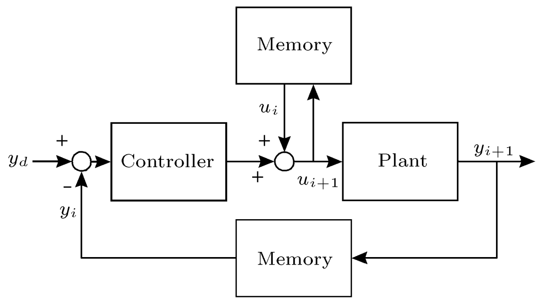

图 2 前馈式学习控制结构

Fig. 2 Configuration of previous cycle learning

2 学习控制两大学派的兴起

经过众多学者的不懈努力，学习控制理论体系日臻完善，现已成为控制方法论中的一重要分支，可与任何主流控制方法，如自适应控制，最优控制，模型预测控制，滑动模块控制等，相颉颃．

要判断一控制理论是否能成为体系需要从多方面考察．每一类控制方法的提出，都是针对某一类受控对象，以及某一类控制目标或任务．例如，自适应控制主要是针对具有未知常系数的线性或非线性受控对象，渐进追踪任给的光滑目标轨道．滑动模块控制则主要是针对具有有界扰动的线性或非线性受控对象，渐进追踪任给的光滑目标轨道．为了避免讨论过于拘泥细节，我们就从这两方面来考察学习控制．

最初的学习控制 – 迭代学习控制，由日本学者有本卓首倡于 1984 年[1]. 前节所述的学习控制律 (2) 为迭代学习控制的最简表达式．不像其他的控制方法从线性受控对象起步，迭代学习控制开门见山就把非线性系统作为研究对象，且要在有限区间 $[0,T]$ 上实现输出完全追踪的控制任务．这里完全追踪 (perfect tracking) 指系统输出自始至终，无论是暂态

或稳态，都要保持和目标轨道一致．显然，迭代学习控制的起点要比其它控制方法高出一截．可是，从二十年的发展历程看，起点过高也有不利的一面：发展空间不足以及难以和主流控制方法相融合．

迭代学习控制从一开始就将目标定于输出完全追踪，避免使用状态变量和状态空间的任何信息．设受控对象为全局 Lipschitz 连续的动力系统

$$\dot{x}(t)=f(x,u,t),\quad y(t)=g(x,u,t) \tag{3}$$

目标轨道为 $y_r(t)\in C^1[0,T]$. 迭代学习控制的任务是要实现 $e(t)=y_r-y(t)$ 在区间 $[0,T]$ 一致收敛于 0. 反馈学习控制律 (1) 或纯前馈式学习控制律 (2) 为迭代学习的两主要算法．学习收敛条件为

$$\left|1-\frac{\partial g}{\partial u}q\right|\leqslant\gamma<1 \tag{4}$$

由 (4) 可以看到迭代学习控制只需要 $\partial g/\partial u$ 的上下界，就可设定学习增益 q 以确保 (4). 和其它控制方法相比，迭代学习控制可称之为几乎无模型控制，完全不需要知道 $f(x,u,t)$ 是何方神圣，稳定还是不稳定．由此可见，迭代学习控制的起点的确很高：用简单得不能再简单的一招学习律，对付具有高度不确定和非线性的动力系统 (3). 可是，这也给想要进一步发扬光大本派的长老弟子们带来了诸多困难．

其实，迭代学习控制难以和主流控制方法相融合的主要原因还不在于起点高，而在于其独特的学习收敛理论核心：压缩映像与不动点原理．迭代学习的主旨是要让输出误差序列按几何级数收敛

$$\|e_i\|\leqslant\gamma\|e_{i-1}\| \tag{5}$$

不胜遗憾的是，当今被非线性控制理论的各大流派捧为宝典的 Lyapunov 方法，却无法渗入经典的迭代学习控制．笔者曾有过不止一次的经历：审稿者表示难以理解迭代学习控制为何不用 $f(x,u,t)$ 的任何信息，为何不用状态量，为何控制系统沿时间轴可以是不稳定的．笔者也曾试图在原有的迭代学习理论框架内作种种改进，但都归于失败．压缩映像与不动点原理原本只在局部域上成立．有本卓引入 $f(x,u,t)$ 为全局 Lipschitz 连续的条件和时间加权模，从而将迭代学习延拓到任意有界区间 $[0,T]$. 很多研究者包括笔者都试图将 $f(x,u,t)$ 放宽为局部 Lipschitz 连续，均铩羽而归．另一方面，非线性控制理论的主流派，如自适应控制与滑动模块控制，运用 $f(x,u,t)$ 的信息，都能轻松自如地对付局部 Lipschitz 连续函数，例如 x^2. 失败与成功的对比，令人悟到，有关 $f(x,u,t)$ 的信息在实现复杂非线性控制中或是不可缺的．沿此方向探索的结果，果然是柳暗花明又一村．在第四节，我们将专门介绍．

就在迭代学习控制被提出的同时，东瀛的控制理论专家们又在酝酿着另一学习控制理论 – 反复控制，又称反复学习控制[2]. 这理论堪称迭代学习控制的孪生兄弟．同样着眼于利用受控对象和目标轨道的周期性或重复性，用前一周期的控制信息去改善后一周期的控制，同样大量地使用存储器以纪录过去的控制信息．只是，迭代学习控制志在有限区间上完全追踪，而反复控制关心的是无限区间上渐近追踪．反复控制律的最简形式为

$$u(t)=u(t-T)+qe(t) \tag{6}$$

我们完全可以将两学习控制律 (1) 与 (6) 在形式上统一起来．首先将无限区间 $[0,\infty]$ 表达为无限多个单位区间 $[(i-1)T,iT]$, $i=1,2,\cdot$. 令 $\tau\in[0,T)$, 则对于 $[0,\infty)$ 上的任一时刻 t,

总能找到一对应区间 $[(i-1)T, iT)$, 使 $t=\tau+(i-1)T$. 记 $u(t)=u_i(\tau)$, $e(t)=e_i(\tau)$, 则 $u(t-T)=u_{i-1}(\tau)$. 于是控制律 (6) 可表达为 $u_i(\tau)=u_{i-1}(\tau)+qe_i(\tau)$, 和 (1) 完全一致. 反复控制的提出，主要是针对伺服 (servo) 控制任务. 根据内模原理，若想要受控对象追踪某一目标轨道或消除某一扰动，控制回路必须内含产生此目标轨道 / 扰动的动力系统模型的全部信息. 让我们考察这样一情况：目标轨道 / 扰动为 C^1 空间的任意周期函数. 根据付利叶级数，以为周期的函数可表达为

$$h(t)=\sum_{j=0}^{\infty} a_j\cos\left(\frac{2\pi j}{T}t\right)+b_j\sin\left(\frac{2\pi j}{T}t\right)$$

系数 a_j 和 b_j 的值取决于周期函数的形状. 显然，产生此目标轨道 / 扰动的动力系统模型很可能为无限维，在拉氏复平面上有无限多对共轭极点位于虚轴上. 要将无限维的动力系统模型的全部信息一一编码写入控制律中显然是不可能的. 何况，我们还不一定能知道这些系数，例如周期扰动. 反复控制方法的精华在于，只要事先知道函数的周期 T, 就能实现控制目标. 反复控制和迭代学习控制一样，并不学习这些周期的目标轨道或扰动，而是直接学习所需的控制量. 其做法是沿学习轴 i 构造一积分器. 这积分器在每一周期进行逐点积分 (pointwise integration), 而每一点的积分运算只和前一周期的同一点发生关系. 让我们考察这逐点积分究竟有何奥妙. 周期函数在相隔周期 iT 的点上值恒同. 因而逐点积分事实上在每一点沿学习轴 i 对付的只是一常数. 对于确定性的受控对象，相应于周期目标或扰动的理想控制也应为一周期函数，即 $u_r(t)=u_r(t-T)=\cdots=u_r(t-iT)=\cdots$. 反复控制正是针对这一类控制问题的妙招.

反复控制理论与方法在伺服 (servo) 控制中得到了广泛的应用. 笔者之一曾在 1996 年作过文献检索，发现有关反复控制的文献数远远高于迭代学习控制. 但是，反复控制理论的发展却不如迭代学习控制来的活跃. 近几年在国际主要控制刊物上发表的学习控制理论一面倒地集中于迭代学习控制. 其原因，笔者认为，在于反复控制的学习收敛条件是在频域中检验的. 固然，反复控制的频域设计方便了实际应用，这也是反复控制的文献数高于迭代学习控制原因. 但另一方面却妨碍了其推广到非线性系统. 一直到最近一、二年，才有了突破. 我们将在第 4 节加以介绍.

3 经典学习控制理论的二十年沧桑

要将学习控制理论二十年的发展成果总结成短短一节文字，几乎是不可能的. 幸好，近几年接连出版了四本有关迭代学习控制的著作，因此在此只选最主要的几项加以介绍.

首先值得一提的是韩国 KAIST 的 Bien 和笔者编著于 1998 年的 ≪ 迭代学习控制 – 分析、设计、综合与应用 ≫[3]. 二十几名学习控制的世界高手，包括本派创始人有本，元老 Bien 、 Moore 、 Longman, 以及众多后起之秀，将他们多年钻研的心得按照分析、设计、综合与应用写成四大部十八章，是对经典学习控制理论从 1984 到 1998 十五年的一次全面总结. 1999 年经典学习控制理论又迎来了另一高潮. 两位弄潮儿陈阳泉，孙明轩和他们的合作者分别出版了关于迭代学习控制的专著. 陈阳泉等的专著[4] 提出、分析和总结了迭代学习控制的一些重要算法与设计. 孙明轩等的专著[5] 则从系统正则性、学习律、初始条件、鲁棒性、分析手段五方面相当系统地回顾了迭代学习控制理论到 1996 年为止所取得的结果. 有意思的是，孙明轩等的专著是以中文发表的. 这标志中国在迭代学习控制方面的研

究到达了当时世界级水平. 笔者之一和博士生谈瑛在 2003 年出版了专著 ≪ 线性与非线性迭代学习控制 ≫[6], 其中前六章集中探讨了经典的迭代学习控制问题与方法. 接下来让我们对当今迭代学习控制理论前沿的几大热点加以考察.

高阶迭代学习律

高阶迭代学习律的一般表达式为

$$u_i(t)=\sum_{j=1}^{n}a_ju_{i-j}(t)+\sum_{j=1}^{n}b_je_{i-j}(t)$$

其出发点非常直观：既然在学习过程只用了前一周期的控制信息就已效果显著，如能利用前几个周期的信息，应更是锦上添花. 高阶迭代学习律由韩国的 Bien 于 1989 年提出[7], 并推导出可加快学习收敛速度的优点. 笔者后来发现其解不成立. 但是文章中提出的概念新颖动人，引众多英雄竞折腰. 在 2002 年 IFAC 世界大会上，学习控制专家们还组织了高阶迭代学习专场. 先前认为高阶迭代可以加速收敛，但据笔者研究结果，若仅限于线性组合过去的控制信息，此论恐难成立[8]. 也有认为高阶迭代可以改善沿学习轴的过渡响应 (如单调收敛等), 但缺少严谨的理论支持. 另有几种观点，有待进一步研究. 一种认为，高阶迭代可以学习沿迭代轴呈规则变化的信号. 又一种认为，高阶迭代可以减少噪声信号的影响. 事实上，有本研究组早在 1987 年就已对高阶迭代进行了考察，得出了频域上改善学习的结论[9]. 只是很遗憾，日文与中文一样未能成为国际学术界的公共语言.

鲁棒迭代学习律

在实际应用中，算法 (1) 或 (2) 均不理想. 由于采样器，算法 (1) 中本周期误差信号被延迟，从而会引起震荡. 算法 (2) 中没有反馈，因而欠鲁棒性. 解决办法是将算法 (1) 和 (2) 合二为一，互相取长补短

$$u_i(t)=u_{i-1}(t)+q_1e_{i-1}(t)+q_2e_i(t) \tag{7}$$

然而，实际应用结果显示 (7) 也不尽如意. 学习误差总是会先降后升，最后不可收拾. 对此，专家们一致认为，学习控制沿迭代轴 i 是一典型的积分器. 实际系统的响应总是包含了可重复与不可重复两大类信号. 另一方面，无论什么信号，学习控制都一积了之. 如果不可重复信号沿迭代轴非 l^2, 就会越积越高. 学习误差会先降，是由于可重复误差的减少. 而后升，是不可重复信号在滚雪球.

怎么解决滚雪球的问题？答案是加滤波器. 于是学习律又被改造成

$$u_i(t)=A(j\omega)u_{i-1}(t)+B(j\omega)e_{i-1}(t)+C(j\omega)e_i(t) \tag{8}$$

其中 $A(j\omega)$, $B(j\omega)$ 和 $C(j\omega)$ 可以是形形色色的滤波器. 经常在实际中使用的最简形式滤波器是加一遗忘因子

$$u_i(t)=\alpha u_{i-1}(t)+q_1e_{i-1}(t)+q_2e_i(t) \tag{9}$$

当系数 $0<\alpha<1$ 时，学习控制沿迭代轴 i 变成一阶低通滤波器. 不过此法仍有缺陷：将有用与无用信号，不分青红皂白统统忘却. 在实际应用中，可重复信号，通常为有用信号，往往位于低频段. 而不可重复信号，如噪声，则位于高频段. 由此生出一两全其美之法：将 $A(j\omega)$ 设计成低通滤波器，保留低频段，衰减高频段. 根据笔者的经验，将 $A(j\omega)$ 设计成二阶 butterworth 滤波器，可以获得良好的效果. 当受控对象为线性时，也可用 H_∞ 等频域的系统设计方法[10], 或是类似状态空间的卡尔曼滤波设计方法[11], 来确定 (8) 中的滤波器.

预测最优迭代学习

对于线性受控对象，模型预测控制和最优控制方法均可推广到线性迭代学习．定义二次型目标函数

$$J_{i+1}(u_{i+1}) = \sum_{j=1}^{N} \lambda^j (\|e_{i+j}\|^2 + \|u_{i+j} - u_{i+j-1}\|^2) \tag{10}$$

N 为预测步长，λ 为权系数．对其求极值可得最优解

$$u_i(t) = u_{i-1}(t) + K_{N-1} e_i(t)$$

K_{N-1} 为最优增益矩阵．相应的最优误差有如下的迭代关系

$$e_i(t) = L_N e_{i-1}(t)$$

L_N 和 K_{N-1}, 随着 N 和 λ 的增加而单调下降．或问，加预测对学习控制有何好处？研究结果显示[12], 预测可以加速学习收敛，尤其是对非最小相位系统，以及改善抗扰性．不过，任何新方法都伴有副作用．此方法需要 $N-1$ 个模型同步运行，计算量大．关键的一点是，预测最优迭代学习的精度高度依赖模型精度．和其它经典学习控制方法相比，这可说既是长处也是短处．长处是能充分利用模型知识，短处是若无完整模型知识则成了一泥足巨人，难有作为．

简单非线性迭代学习律

一般非线性迭代学习律本身可以是一动力系统，表达为

$$\begin{aligned} \dot{x}_c(t) &= f_c(x_c, e_i, \cdots, e_{i-n}, t) \\ u_i(t) &= g_c(x_c, u_{i-1}, \cdots, u_{i-n}, e_i, \cdots, e_{i-n}, t) \end{aligned} \tag{11}$$

显然，前述的各类迭代控制均为 (11) 的特殊形式．尽管非线性迭代学习律已被提出很久了，一直未见有系统性的设计方法问世．迭代学习只须 $f_c(\cdot)$ 和 $g_c(\cdot)$ 是全局 Lipschitz 连续，以及知道 $\partial g_c(\cdot)/\partial u$ 的上下界，就能完成重任．不过，一个令人惊讶事实是，我们并没有设计 $f_c(\cdot)$ 和 $g_c(\cdot)$ 的妙法．如果不能证明非线性学习控制优于线性学习控制，顾客是不会光顾非线性产品的．笔者的研究组在参考了数值分析方法的基础上提出了两种非线性迭代学习律[13,6]：牛顿学习律 (12) 和正割 (Secant) 学习律 (13) 分别如下

$$u_i(t) = u_{i-1}(t) + q\frac{e_{i-1}(t)}{\partial g/\partial u} \tag{12}$$

$$u_i(t) = u_{i-1}(t) + q\frac{e_{i-1}(t) - e_{i-2}(t)}{u_{i-1} - u_{i-2}} \tag{13}$$

可以看出，牛顿学习律需要用到系统输出方程的偏导，而不仅仅是偏导的上下界．正割学习律事实上构成了一简单的二阶非线性迭代．根据理论分析，牛顿学习律和正割学习律均能加快学习收敛速度，相对收敛速度关系为：牛顿学习律 $>$ 正割学习律 $>$ 线性学习律．代价是需要用到更多的系统信息．

4 新学习控制理论的登场

前面提到了学习控制理论的柳暗花明，峰回路转．在此介绍柳暗花明村新成员．需要强调的是，经典学习控制研究正方兴未艾，并非古道夕阳．只是，在构造经典学习控制器

时，我们追求是的压缩映射这一数学上完美的形式，因为压缩映射的构造同时还保证了解的存在性与唯一性，以及几何收敛速度．但是压缩映射偏重系统输入与输出关系，很少能充分利用系统状态信息，导致经典学习控制的设计方法只能局限于全局 Lipschitz 连续的非线性系统．

让我们来构造一般的非线性学习控制．在此，我们将放弃压缩映射这一完美却无法实现的目标．事实上，对于林林总总的非线性动力系统，只要能证明解的存在和控制的渐进收敛，我们就非常的心满意足了．的确，基于 Lyapunov 的主流非线性控制方法，如自适应控制、非线性鲁棒控制等，莫不如此．新学习控制理论的第一要务，是解决局部 Lipschitz 连续的动力系统，例如微分方程右边含有 x^2 项．其次，还需考虑不确定要素，包括参数型与非参数型．为突出概念，考察最简单的受控对象

$$\dot{x}(t)=\eta(x,t)+u(t) \tag{14}$$

若 $\eta(x,t)$ 可分解为 $a(t)x^2$ 且 $a(t)$ 为一未知、与状态 x 无关的参数，而非线性项 x^2 已知，称系统 (14) 具有参数型不确定要素．如果 $\eta(x,t)$ 无法分解为未知参数和已知状态函数，如 $\eta(x,t)=x^2\sin x$ 整体未知，称系统 (14) 具有非参数型不确定要素．自适应控制与非线性鲁棒控制这两大派的主旨就是要分别对付参数型或非参数型不确定要素的控制问题．我们当然希望学习控制也能对付这两类不确定要素，并且要青出于蓝而胜于蓝．新学习控制理论还要能继承前辈们的优良传统：迭代学习控制的有限区间完全追踪和反复控制的周期性学习．

综合上述的期望，新学习控制理论必须是一开放理论体系，最好能兼容各家学派，百川汇海．作为非线性控制理论，还要有一行之有效的稳定性分析方法．众所周知，基于 Lyapunov 函数以及 Lyapunov 泛函的稳定性分析是现代非线性控制理论的镇山之宝．因此，能否运用这一镇山之宝，将是区别经典与新学习控制理论的主要标志之一．

打响第一枪的是来自佛罗里达中央大学的瞿之华研究组．在他们学会论文中[14]，讨论了有限区间上参数型系统的学习问题，引入了泛函

$$V(\hat{a}_i,t)=\int_0^t[a(t)-\hat{a}_i(\tau)]^2\mathrm{d}\tau \tag{15}$$

和参数型迭代学习律

$$\hat{a}_i(t)=\hat{a}_{i-1}(t)+qx_i^2e_i(t),\quad t\in[0,T] \tag{16}$$

其中 $e_i(t)=x_r(t)-x_i(t)$ 是状态追踪误差．值得注意的是，非线性学习律 (16) 终于用上了动力系统的局部 Lipschitz 项 x^2．有意思的是，该研究组长期从事非线性鲁棒控制的研究，因而能跳出经典学习的框架，用鲁棒之石，攻学习之玉．

我们再来看一下，学习控制可否从自适应控制中也学几手妙招．英国的 Rogers 研究组在探讨有限区间上参数型系统的学习问题时[15]，引入 Lyapunov 函数 $V(e_i,\hat{a}_i,t)$

$$V(e_i,\hat{a}_i,t)=e_i^2(t)+[a-\hat{a}_i(t)]^2$$

和微分型参数学习律

$$\dot{\hat{a}}_i(t)=qx_i^2(t)e_i(t),\quad t\in[0,T]$$

如果不考虑迭代指标 i, 以上的微分型参数学习律和自适应控制一般无二. 可是在有限区间上我们却享受不到自适应控制沿时间轴渐近收敛的长处. 解决方法是将相邻两次参数迭代学习通过初、终值挂钩

$$\hat{a}_i(0) = \hat{a}_{i-1}(T)$$

此法被称为自适应学习控制, 可看作是自适应控制到学习控制之间的过渡, 因为只适用于未知定常参数的场合. 而方法 (16) 则适用于任意连续时变参数.

经过多年的努力, 新学习控制理论的框架已初步形成[6,16~27]. 接下来, 以 Lyapunov 泛函为中轴, 简要介绍四类典型学习控制任务.

有限区间参数型学习

设受控对象为

$$\dot{x}(t) = a(t)x^2(t) + u(t)$$

目标轨道为 $x_r(t)$, $t \in [0, T]$. 控制律和参数学习律分别为

$$u_i(t) = \dot{x}_r(t) + ke_i(t) + \hat{a}_i(t)x_i^2(t), \quad \hat{a}_i(t) = \hat{a}_{i-1}(t) + qx_i^2(t)e_i(t) \tag{17}$$

定义 Lyapunov 泛函

$$V(e_i, \hat{a}_i, t) = \frac{1}{2}e_i^2(t) + \frac{1}{2q}\int_0^t [a(t) - \hat{a}_i(\tau)]^2 \mathrm{d}\tau$$

可以推导出

$$\Delta V_i = V(e_i, \hat{a}_i, t) - V(e_{i-1}, \hat{a}_{i-1}, t) \leqslant -\int_0^t e_i^2(\tau)\mathrm{d}\tau$$

进而推导出

$$\lim_{i\to\infty}\int_0^{\mathrm{T}} e_i^2(\tau)\mathrm{d}\tau = 0$$

由此得出：系统状态 $x_i(t)$ 在区间 $[0, T]$ 上几乎处处收敛于目标轨道 $x_r(t)$.

无限区间参数型学习

设受控对象为

$$\dot{x}(t) = a(t)x^2(t) + u(t)$$

目标轨道为 $x_r(t)$, $t \in [0, \infty)$. $a(t) = a(t - T)$ 是以 T 为周期的未知时变参数. 控制律和参数学习律分别为

$$u(t) = \dot{x}_r(t) + \hat{a}(t)x^2(t), \quad \hat{a}(t) = \hat{a}(t - T) + qx^2(t)e(t) \tag{18}$$

定义 Lyapunov 泛函

$$V(e, \hat{a}, t) = \frac{1}{2}e^2(t) + \frac{1}{2q}\int_{t-T}^t [a(t) - \hat{a}(\tau)]^2 \mathrm{d}\tau$$

可以推导出 Lyapunov 泛函的右上导数

$$\dot{V} \leqslant -\int_{t-T}^t e^2(\tau)\mathrm{d}\tau$$

进而推导出

$$\lim_{t\to\infty}\int_{t-T}^{t} e^2(\tau)\mathrm{d}\tau = 0$$

由此得出：系统状态 $x(t)$ 渐进地几乎处处收敛于目标轨道 $x_r(t)$.

有限区间非参数型学习

设受控对象为

$$\dot{x}(t) = x^2(t)\sin x(t) + u(t)$$

目标轨道为 $x_r(t)$, $t\in[0,T]$. 如果我们对非线性不确定项 $x^2(t)\sin x(t)$ 一无所知，任何控制方法都将无能为力，学习也将无从谈起. 在实现非线性鲁棒控制或滑动模块控制中，需要知道非线性不确定项的上界函数 $\beta(x,t)\geqslant|x^2(t)\sin x(t)|$. 为平等起见，在实现学习控制时，也设 $\beta(x,t)$ 已知.

此时此刻，鲁棒派长老们也许会说，既然本派已能保证追踪精度，学习控制不用也罢. 然而, 其中有很大的区别. 鲁棒控制本质上是高增益反馈. 追踪精度取决于增益高低. 理论上讲，要实现零误差需无限高增益. 这是无法实现的. 非但如此，在实际控制中由于采样的关系，增益往往稍高一点都不行. 当系统重复可学时，我们多了一个选择：先构造一鲁棒控制器，保证有界追踪误差. 如果增益不能取得很高，误差可能会很大. 不过不要紧，学习控制正等着一显身手. 通过反复学习，可以将有界误差逐渐减到零. 从而我们有了一新控制方法：鲁棒学习控制，综合了两大学派的长处.

下面给出鲁棒学习律的一特定形式[26]

$$\begin{aligned} &u_i(t) = proj[u_{i-1},u^*] + v_i,\ v_i(t) = (\rho_i\kappa_i + 1)e_i \\ &\rho_i = \sqrt{|\dot{x}_r|^2+\varepsilon+\beta_i},\ \kappa_i = \frac{\sqrt{e_i^2+3\varepsilon^2}+8\varepsilon}{(\sqrt{e_i^2+3\varepsilon^2}+\varepsilon)^2} \end{aligned} \tag{19}$$

在此 $\beta_i=\beta(x_i,t)$, $proj[u_{i-1},u^*]$ 为一限幅器，u^* 为理想控制信号 $u_r(t)$ 上限的一充分大估计值. v_i 为鲁棒控制，若 $u_i=v_i$, 可以推出其误差界为 ε.

定义 Lyapunov 泛函

$$V(e_i,t) = \frac{1}{2}e_i^2(t) + \frac{1}{2}\int_0^t e^{-\lambda\tau}\delta u_i^2(\tau)\mathrm{d}\tau \tag{20}$$

可以推导出

$$\Delta V_i = V(e_i,t) - V(e_{i-1},t) \leqslant -e^{-\lambda t}e_{i-1}^2$$

进而推导出

$$\lim_{i\to\infty} e_i^2(t) = 0$$

系统状态 $x_i(t)$ 在区间 $[0,T]$ 上处处收敛于目标轨道 $x_r(t)$. 鲁棒学习控制可以实现低增益下的完全追踪.

无限区间非参数型学习

最后考察在无限区间非参数型学习. 在此，目标轨道 $x_r(t)$ 是以 T 为周期的函数. 控制器结构可以设计得和有限区间的鲁棒学习控制 (19) 几乎完全一样. 唯一的不同之处在于鲁棒控制部分. 我们注意到 Lyapunov 泛函 (20) 中引入了时间忘却因子 $e^{-\lambda t}$, 参数 λ 可

以是一任意大的数．该因子引入，是要抑制在紧集上非线性不确定项 $\eta(x,t)$ 的 Lipschitz 系数．但此法显然不适合于无限区间．因此有必要修正 (17) 的鲁棒控制部分

$$u(t)=proj[u(t-T),u^*]+v,\ v(t)=(\rho\kappa+1)e+\gamma$$
$$\rho=\sqrt{|\dot{x}_r|^2+\varepsilon}+\beta,\ \kappa=\frac{\sqrt{e^2+3\varepsilon^2}+8\varepsilon}{(\sqrt{e^2+3\varepsilon^2}+\varepsilon)^2} \tag{21}$$

其中 $\gamma=\gamma(x,t)$ 为 $\eta(x,t)=x^2\sin x$ 的偏导数的上界函数．可以证明系统状态 $x(t)$ 渐进地几乎处处收敛于目标轨道 $x_r(t)$ [27]．

5 学习控制理论的今天与明天

经历了二十年的风风雨雨，经典学习控制理论已是树大根深，枝繁叶茂．凭其独特的学习能力，以及和 PID 一样简洁的算法，学习控制正在应用领域攻城拔寨．无论是控制目标、受控对象还是外界扰动，只要具有周期性或可重复性，学习控制总能胜任主角或主要配角，大幅度提高控制精度．和其它的学习方法，如神经元网络学习，统计学习相比，经典的迭代学习和反复学习着眼于单一目标的精确学习，任务明确单纯，从而学习算法可以大为简化．最简单的学习控制律 (1) 只有一比例反馈加上若干记忆存储器．结构简单、针对性强、效果显著这三个经典学习控制的特点正是实际应用所强调的．最初的迭代学习控制往往以机器人手臂为控制对象．最近的研究报告显示，迭代学习的概念被尝试应用于过程控制[28,29]，电气驱动[30,31]，精密伺服[32]，以及系统建模[32,33] 等．反复控制则广泛地被试用于各种伺服系统．

另一方面，以 Lyapunov 泛函为核心的新学习控制理论正忙着在理论领域里攻城略地，要将学习控制发展为一完整的控制理论体系．对此，自适应控制理论体系堪称是良师楷模．关于系统相对阶数、基于状态或输出信号的反馈，参考模型与系统之间的制约，自适应观察器的构造，多种鲁棒化修正律，解的连续性和存在性，以及控制器设计等的深入探讨，标志着自适应控制已形成一完整理论体系[34,35]．相比之下，新学习控制理论才刚开始起步．同时，针对上述的系统控制的核心问题，新学习控制理论要解决的问题难度则远高于自适应控制．以无限区间参数型学习过程的解的存在性为例．系统和学习控制律 (18) 构成一混合型微分和连续时间差分方程组．因而无法像自适应控制理论那样套用传统的基于微分方程的解的存在性理论．考虑到自适应控制理论体系的建立花了几代高手的近四十年大好年华，学习控制理论的体系化任重而道远，更有待于各路英雄豪杰一展身手，夺关攻擂．

让我们从较为数学的角度来考察学习控制理论，探讨新的研究方向．经典学习控制事实上在 Banach 空间中构造了一类特殊的压缩映射，生成一基本点列，因而学习收敛的极限就是唯一的不动点．由于压缩映射的构造颇为不易，因此另一可能的途径是引用其他的不动点理论，构造新学习算子，扩大可学习的范围．此外，众所周知许多控制方法只在内积空间中有效．如何在内积空间对经典学习控制方法进行加工改造，发挥内积空间的特点，有待进一步研究．

再考察新学习控制理论的发展空间．这里仅举三点．首先，第四节柳暗花明村中介绍的学习控制的本质是函数逼近：逼近未知函数 $a(t)$ 或 $u_r(t)$．在泛函中对函数逼近有众多精彩的理论结果．如何开发这一金矿，应用到闭环控制系统中，例如将函数逼近与 Lyapunov

泛函结合，是学习控制理论专家们的重任之一．事实上，神经元网络和小波网络被控制理论看中的也正是其独特的函数逼近能力．这方面的先驱工作有基于小波网络的学习控制[36,37]．其次，仅考虑单一周期的学习是不够的．有必要将学习延拓到拟周期或概周期，甚至非周期函数．最后，作为开放的理论体系，基于 Lyapunov 泛函的学习控制还应当是"拿来主义"的高手，能吸取各大控制理论流派的精华，产生新一代如非线性最优学习控制，非线性模型预测学习控制，模糊学习控制，统计学习控制，生物学习控制等等．

References

1 Arimoto S, Kawamura S, Miyazaki F. Better operation of robots by learning. *Journal of Robotic Systems*, 1984, **1**(2): 123~140

2 Hara S, Yamamoto Y, Omata T, Nakano M. Repetitive control systems: a new type servo system for periodic exogenous signals. *IEEE Transactions on Automatic Control*, 1988, **33**(7): 659~668

3 Bien Z, Xu J-X. Iterative Learning Control – Analysis, Design, Integration and Applications. Boston: Kluwer Academic Press, 1998

4 Chen Y Q, Wen C Y. Iterative Learning Control - Convergence, Robustness and Applications. In series of Lecture Notes in Control and Information Sciences 248. Berlin: Springer Verlag, 2003

5 Sun M X, Huang B J. Iterative Learning Control. Beijing: Defence Industry Press, 1999

6 Xu J-X, Tan Y. Linear and Nonlinear Iterative Learning Control. In series of Lecture Notes in Control and Information Sciences 291, Berlin: Springer Verlag, 2003

7 Bien Z, Hui K M. High-order iterative learning control algorithm. *IEE Proceedings, Part D, Control Theory and Applications*, **136**(3): 105~112

8 Xu J-X, Tan Y. Robust optimal design and convergence properties analysis of iterative learning control approaches. *Automatica*, 2002, **38**(11): 1687~1880

9 Gotou M, Matsubayashi S, Miyazaki F, Kawamura S,Arimoto S. A robust servo system with an iterative learning compensator and a proposal of multi-period learning compensator. *System and Control Magazine*, 1987, **31**(5): 367~374

10 Roover D. Synthesis of a robust iterative learning control system using an H_∞ approach. In: Proceeding of 35th IEEE Conference on Decision and Control, Kobe, 1996. 3044~3049

11 Saab S S. A discrete-time stochastic learning control algorithm. *IEEE Transactions on Automatic Control*, 2001, **46**(6): 877~887

12 Amann N, Owens D H, Rogers E. Predictive optimal iterative learning control. *International Journal of Control*, 1998, **69**(2): 203~226

13 Xu J-X, Tan Y. On the P-type and Newton-type ILC schemes for dynamic systems with non-affine-in-input factors, *Automatica*, 2002, **38**(7): 1237~1242

14 Ham C, Qu Z H, Kaloust J H. Nonlinear learning control for a class of nonlinear systems based on Lyapunov′s direct method. In: Proceeding of 1995 IEEE American Control Conference, Seattle, 1995. 3024~3028

15 Frueh M, Rogers E. Nonlinear iterative learning by an adaptive Lyapunov technique. *International Journal of Control*, 2000, **73**(10): 840~850

16 Xu J-X, Qu Z H. Robust iterative learning control for a class of nonlinear systems. *Automatica*, 1998, **34**(8): 983~988

17 Xu J-X, Badrinath V. Adaptive Robust Iterative Learning Control with Dead Zone Scheme. *Automatica*, 2000, **36**(1): 91~99

18 Xu J-X, Badrinath V, Qu Z H. Robust learning control for robotic manipulators with an extension to a class of non-linear systems. *International Journal of Control*, 2000, **73**(10): 858~870

19 Xu J-X, Tan Y. A suboptimal learning control scheme for nonlinear systems with time-varying parametric uncertainties. *Journal of Optimal Control – Applications and Methods*, 2001, **22**(4): 111~126

20 Xu J-X, Cao W J. Learning variable structure control approaches for repeatable tracking control tasks. *Automatica*, 2001, **37**(7): 997~1006

21 Qu Z H, Xu J-X. Model based learning control and their comparisons using Lyapunov direct method. *Asian Journal of Control*, 2002, **4**(1): 99~110

22 Cao W J, Xu J-X. On Functional approximation of the equivalent control using learning variable structure control. *IEEE Transactions on Automatic Control*, 2002, **47**(5): 824∼830
23 Xu J-X, Tan Y. A composite energy function based learning control approach for nonlinear systems with time-varying parametric uncertainties. *IEEE Transactions on Automatic Control*, 2002, **47**(11): 1940∼1945
24 Xu J-X. The frontiers of iterative learning control – Part I, *Journal of Systems, Control and Information*, 2002, **45**(2): 63∼73
25 Xu J-X. The frontiers of iterative learning control – Part II, *Journal of Systems, Control and Information*, 2002, **45**(5): 233∼243
26 Xu J, Xu J-X. Memory-based nonlinear internal model – what can a control system learn. In: Proceedings of 4th Asia Control Conference, Singapore, 2002. 446∼451
27 Tan Y, Xu J X. Learning based nonlinear internal model control. In: Proceedings of IEEE 2003 American Control Conference, Denver, 2003. 3009∼3013
28 Lee K S, Lee J H. Model-based predictive control combined with iterative learning for batch or repetitive processes. Chapter 16, in: Iterative Learning Control – Analysis, Design, Integration and Applications, edited by Bien Z. and Xu J-X, Boston, USA: Kluwer Academic Press, 1998
29 Xu J-X, Hu Q P, Lee T H, S Yamamoto S. Iterative learning control with Smith time delay compensator for batch processes. *Journal of Process Control*, 2001, **11**(3): 321∼328
30 Sahoo S K, Panda S K, Xu J-X. Iterative learning based high performance current controller for switched reluctance motor. *IEEE Transactions on Energy Conversion*, 2004, **19**(3): 491∼498
31 Qian W Z, Panda S K, Xu J-X. Torque ripples minimization in PM synchronous motors using iterative learning control, *IEEE Transactions on Power Electronics*, 2004, **19**(2): 272∼279
32 Xu J-X, Lee T H. Development of nonlinear precision control for nano tele-manipulation. Technical Report WBS R-263-000-099-112, 2003
33 Chen Y Q, Xu J X, Wen C Y. Iterative learning based extraction of aerobomb drag. *AIAA: Journal of Spacecraft and Rockets*, 1998, **35**(1): 237∼240
34 Narendra K S, Annaswamy A M. Stable Adaptive Control. London: Prentice Hall, 1989
35 Ioannou P A, Sun J. Robust Adaptive Control. London: Prentice Hall, 1996
36 Fukuda M, Shin S. Model reference learning control with a wavelet network. Chapter 11, in Iterative Learning Control – Analysis, Design, Integration and Applications, edited by Bien Z. and Xu J-X, Boston: Kluwer Academic Press, 1998
37 Xu J-X, Tan Y. Learning wavelet control using constructive wavelet networks. Chapter 9, in Linear and Nonlinear Iterative Learning Control, Berlin: Springer Verlag, 2003

许建新 1982 年毕业于浙江大学电机系获本科学位，1989 年毕业于东京大学电气工学系，获得工学博士．1991 年至今在新加坡国立大学电气与计算机系任副教授．1989 至 1990 年在日立公司研究所研修，1990 年至 1991 年在美国俄亥俄州立大学任访问学者，2003 年为美国耶鲁大学访问研究员．研究方向包括：学习控制、变结构控制、模糊控制、不连续信号处理，以及机电系统和化工过程中的控制应用问题．现为 IEEE 高级会员．

(**XU Jian-Xin** Received his bachelor degree from Zhejiang University, P. R. China in 1982. He attended the University of Tokyo, Japan, where he received his master and Ph. D. degrees in 1986 and 1989, respectively. All his degrees are in electrical engineering. He worked for one year in the Hitachi Research Laboratory, Japan; for more than one year in Ohio State University, U.S.A. as a visiting scholar; and for 6 months in Yale University as a visiting research fellow. In 1991 he joined the National University of Singapore, and is currently an associate professor in the Department of Electrical Engineering. His research interests include learning control, variable structure control, fuzzy logic control, discontinuous signal processing, and applications to motion control and process control.)

侯忠生 北京交通大学电子信息工程学院先进控制系统研究所教授．研究兴趣包括非模型自适应控制，学习控制和智能控制．

(**HOU Zhong-Sheng** Professor in the Advanced Control Systems Laboratory of the School of Electronics and Information Engineering, Beijing Jiaotong University. His research interests include the model free adaptive control, learning control, and intelligent control.)