

短文

# 任意初始状态下的学习控制<sup>1)</sup>

任雪梅 高为炳

(北京航空航天大学第七研究室 北京 100083)

## 摘 要

本文对线性时不变系统和一类非线性系统提出了一种新的学习控制方案,此方案包括两种学习,即同时进行系统的输入和初始状态的学习。因此在学习开始时对系统的初始状态无要求。仿真结果表明了这种方法的可行性。

**关键词** 线性系统,非线性系统,学习控制,初始状态。

## 1 引言

近几年来,学习控制<sup>[1-6]</sup>越来越受到人们的注意。它的主要特点是不需要动态系统的精确描述,能够根据以前的操作数据,自动地调节不理想的输入信号,产生一个新的控制,从而使系统的性能得以提高。但对系统的初始状态有限制,即要求在每次循环过程中系统的初始状态  $x_i(0)$  ( $i = 1, 2, \dots, i$  为循环次数)和期望跟踪轨迹的初始状态  $x_d(0)$  都相等<sup>[1,2]</sup>。在实际应用中不易求得期望跟踪轨迹的初始状态,为了消除由于系统的初始状态和期望跟踪轨迹的初始状态的不同而引起的系统输出不能完全跟踪期望轨迹<sup>[6]</sup>,作者提出了同时进行系统的输入和初始状态的学习来实现系统的完全跟踪。

## 2 线性时不变系统中的学习控制

现在考虑下列线性时不变系统:

$$\dot{x}(t) = Ax(t) + Bu(t), \quad (2.1 a)$$

$$y(t) = Cx(t), \quad x(t_0) = x_0. \quad (2.1 b)$$

其中  $x(t) \in R^n, u(t), y(t) \in R^m, A, B, C$  为相应维数的常阵且满足假设

$$\text{rank} CB = m.$$

设系统(2.1)所要跟踪的期望轨迹为  $y_d(t), t \in [0, T]$ 。系统第  $i$  次循环中的输出为  $y_i(t)$ , 令  $e_i(t) = y_d(t) - y_i(t)$ 。在开始学习时对于任意的连续输入  $u_0(t)$  和任意的初始状态  $x_0(0)$ , 现在的问题是已知第  $i$  次运动中的  $u_i(t), x_i(0)$  和  $e_i(t)$ , 怎样求得  $u_{i+1}(t)$  和  $x_{i+1}(0)$ , 使第  $i+1$  次运动误差  $e_{i+1}(t)$  减小。

本文于1992年5月21日收到。

1) 国家自然科学基金资助项目。

为了分析学习控制过程, 首先介绍下列范数.

$$\|e(t)\|_{\infty} = \max_{1 \leq i \leq m} |e^{(i)}(t)|, \quad \|G\|_{\infty} = \max_{1 \leq i \leq m} \left\{ \sum_{j=1}^m |g^{(i,j)}| \right\},$$

$$\|e(t)\|_{\lambda} = \sup_{0 \leq t \leq T} \{ \exp(-\lambda t) \|e(t)\|_{\infty} \}.$$

其中  $e^{(i)}(t)$  是  $e(t) \in R^m$  中的第  $i$  个元素,  $g^{(i,j)}$  是  $G \in R^{m \times m}$  中的第  $i, j$  元素,  $\lambda > 0$ .

系统(2.1)学习控制规律和初始状态学习规律为

$$u_{i+1}(t) = u_i(t) + L\dot{e}_i(t), \quad (2.2)$$

$$x_{i+1}(0) = x_i(0) + BL e_i(0), \quad (2.3)$$

其中  $L \in R^{m \times m}$  是常阵. 于是有

**定理 1.** 若学习控制系统(2.1)–(2.3)满足条件

1)  $u_0(t)$  在  $[0, T]$  上连续,  $y_d(t)$  在  $[0, T]$  上连续可微,

2)  $\|I_m - CBL\|_{\infty} < 1$ ,

则当  $i \rightarrow \infty$  时, 有

$$y_i(t) \rightarrow y_d(t), \quad t \in [0, T].$$

证明. 由(2.1)–(2.3)式得

$$\begin{aligned} y_d(t) - y_{i+1}(t) &= y_d(t) - [C \exp(At)x_i(0) + C \exp(At)BL e_i(0) \\ &\quad + \int_0^t C \exp(A(t-\tau))B u_i(\tau) d\tau \\ &\quad + \int_0^t C \exp(A(t-\tau))BL \dot{e}_i(\tau) d\tau]. \end{aligned} \quad (2.4)$$

因为

$$\begin{aligned} \int_0^t C \exp(A(t-\tau))BL \dot{e}_i(\tau) d\tau &= CBL e_i(t) - C \exp(At)BL e_i(0) \\ &\quad + \int_0^t CA \exp(A(t-\tau))BL e_i(\tau) d\tau, \end{aligned} \quad (2.5)$$

所以(2.4)式可变为

$$e_{i+1}(t) = (I_m - CBL)e_i(t) - \int_0^t CA \exp(A(t-\tau))BL e_i(\tau) d\tau. \quad (2.6)$$

方程两边同乘以  $\exp(-\lambda t)$  并取范数得

$$\begin{aligned} \|e_{i+1}(t)\|_{\lambda} &\leq \|I_m - CBL\|_{\infty} \|e_i(t)\|_{\lambda} \\ &\quad + \sup_{0 \leq t \leq T} \|CA \exp(At)BL\|_{\infty} \int_0^t \exp(-\lambda(t-\tau)) \|e_i(\tau)\|_{\lambda} d\tau \\ &\leq \left[ \|I_m - CBL\|_{\infty} + \frac{1 - \exp(-\lambda T)}{\lambda} \sup_{0 \leq t \leq T} \|CA \exp(At)BL\|_{\infty} \right] \\ &\quad \times \|e_i(t)\|_{\lambda}, \end{aligned} \quad (2.7)$$

故当  $\lambda$  足够大, 由条件(2)可知

$$\rho = \|I_m - CBL\|_{\infty} + \frac{1 - \exp(-\lambda T)}{\lambda} \sup_{0 \leq t \leq T} \|CA \exp(At)BL\|_{\infty} < 1, \quad (2.8)$$

与(2.7)式联立得,当  $i \rightarrow \infty$  时,有

$$\|e_i(t)\|_\lambda \rightarrow 0, t \in [0, T].$$

定理证毕.

由定理 1 的证明过程可知,当  $L = (CB)^{-1}$  时,收敛速度最快,下面讨论此时系统输出误差和速度误差的大小.由(2.6)式得

$$e_{i+1}(t) = - \int_0^t CA \exp(A(t-\tau)) BL e_i(\tau) d\tau, \quad (2.9)$$

$$\dot{e}_{i+1}(t) = - \int_0^t CA^2 \exp(A(t-\tau)) BL e_i(\tau) d\tau - CAB L e_i(t). \quad (2.10)$$

依照证明定理 1 的同样步骤,可得

$$\|e_{i+1}(t)\|_\infty \leq \frac{\alpha^{i+1} t^{i+1}}{(i+1)!} \sup_{0 \leq \tau \leq T} \|e_0(t)\|_\infty, \quad (2.11)$$

$$\|\dot{e}_{i+1}(t)\|_\infty \leq \frac{\alpha^i t^i}{i!} \left( n + \frac{mt}{i+1} \right) \sup_{0 \leq \tau \leq T} \|e_0(t)\|_\infty. \quad (2.12)$$

其中  $\alpha = \sup_{0 \leq t \leq T} \|CA \exp(At) BL\|_\infty$ ,  $m = \sup_{0 \leq t \leq T} \|CA^2 \exp(At) BL\|_\infty$ ,  $n = \|CABL\|_\infty$ .

(2.11)、(2.12) 式给出了当  $L = (CB)^{-1}$  时系统的输出误差和速度误差一致有界的定量范围.

### 3 非线性系统的学习控制

考虑下列形式的非线性系统:

$$\dot{x}(t) = f(x, t) + Bu(t), \quad (3.1a)$$

$$y(t) = Cx(t), x(t_0) = x_0. \quad (3.1b)$$

其中  $x(t) \in R^n$ ,  $u(t), y(t) \in R^m$ , 假设函数  $f(x, t): R^n \times [0, T] \rightarrow R^n$  在  $t \in [0, T]$  上对  $x$  满足 Lipschitz 条件,即

$$\|f(x_1, t) - f(x_2, t)\|_\infty \leq \alpha(t) \|x_1 - x_2\|_\infty.$$

其中  $\alpha(t)$  在  $t \in [0, T]$  上是分段连续函数.

**定理 2.** 若非线性系统(3.1)的学习控制规律和初始状态学习规律为(2.2)、(2.3),且满足定理 1 中的两个条件,则当  $i \rightarrow \infty$  时,有

$$y_i(t) \rightarrow y_d(t), t \in [0, T].$$

证明.由(3.1),(2.2),(2.3)式得

$$e_{i+1}(t) = e_i(t) - C[x_{i+1}(t) - x_i(t)], \quad (3.2)$$

$$x_{i+1}(t) - x_i(t) = BL e_i(t) + \int_0^t [f(x_{i+1}(\tau), \tau) - f(x_i(\tau), \tau)] d\tau, \quad (3.3)$$

方程(3.3)两边同乘以  $\exp(-\lambda t)$ , 取范数得

$$\begin{aligned} \|x_{i+1}(t) - x_i(t)\|_\lambda &\leq \|BL\|_\infty \|e_i(t)\|_\lambda \\ &+ \alpha_0 \int_0^t \exp(-\lambda(t-\tau)) \|x_{i+1}(\tau) - x_i(\tau)\|_\lambda d\tau. \end{aligned} \quad (3.4)$$

其中  $\alpha_0 = \max_{0 \leq t \leq T} \alpha(t)$ .

应用 Bellman-Gronwall Lemma 3.4 式可变为

$$\begin{aligned} \|x_{i+1}(t) - x_i(t)\|_\lambda &\leq \|BL\|_\infty \|e_i(t)\|_\lambda \exp \left\{ \alpha_0 \int_0^t \exp(-\lambda(t-\tau)) d\tau \right\} \\ &\leq \|BL\|_\infty \exp \left\{ \alpha_0 \frac{1 - \exp(-\lambda T)}{\lambda} \right\} \|e_i(t)\|_\lambda. \end{aligned} \quad (3.5)$$

综合(3.2),(3.3)和(3.5)可得定理的证明.

## 4 仿真结果

考虑如下的多输入多输出系统:

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} -2 & 3 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix}, \quad (4.1 a)$$

$$\begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}. \quad (4.1 b)$$

期望跟踪轨迹为

$$\begin{bmatrix} y_{1d}(t) \\ y_{2d}(t) \end{bmatrix} = \begin{bmatrix} 1.5 t \\ 1.5 t \end{bmatrix}, \quad t \in [0, 1].$$

由于  $CB = \begin{bmatrix} 2 & 2 \\ 0 & 1 \end{bmatrix}$ , 故取  $L = \begin{bmatrix} 0.4 & -0.5 \\ 0 & 0.5 \end{bmatrix}$  可满足定理 1 中的条件 (2). 系统

(4.1)的学习控制规律和初始状态学习规律分别为

$$\begin{aligned} \begin{bmatrix} u_{1(i+1)}(t) \\ u_{2(i+1)}(t) \end{bmatrix} &= \begin{bmatrix} u_{1(i)}(t) \\ u_{2(i)}(t) \end{bmatrix} + \begin{bmatrix} 0.4 & -0.5 \\ 0 & 0.5 \end{bmatrix} \begin{bmatrix} e_{1(i)}(t) \\ e_{2(i)}(t) \end{bmatrix} \\ \begin{bmatrix} x_{1(i+1)}(0) \\ x_{2(i+1)}(0) \end{bmatrix} &= \begin{bmatrix} x_{1(i)}(0) \\ x_{2(i)}(0) \end{bmatrix} + \begin{bmatrix} 0.4 & 0 \\ 0 & 0.5 \end{bmatrix} \begin{bmatrix} e_{1(i)}(0) \\ e_{2(i)}(0) \end{bmatrix}. \end{aligned}$$

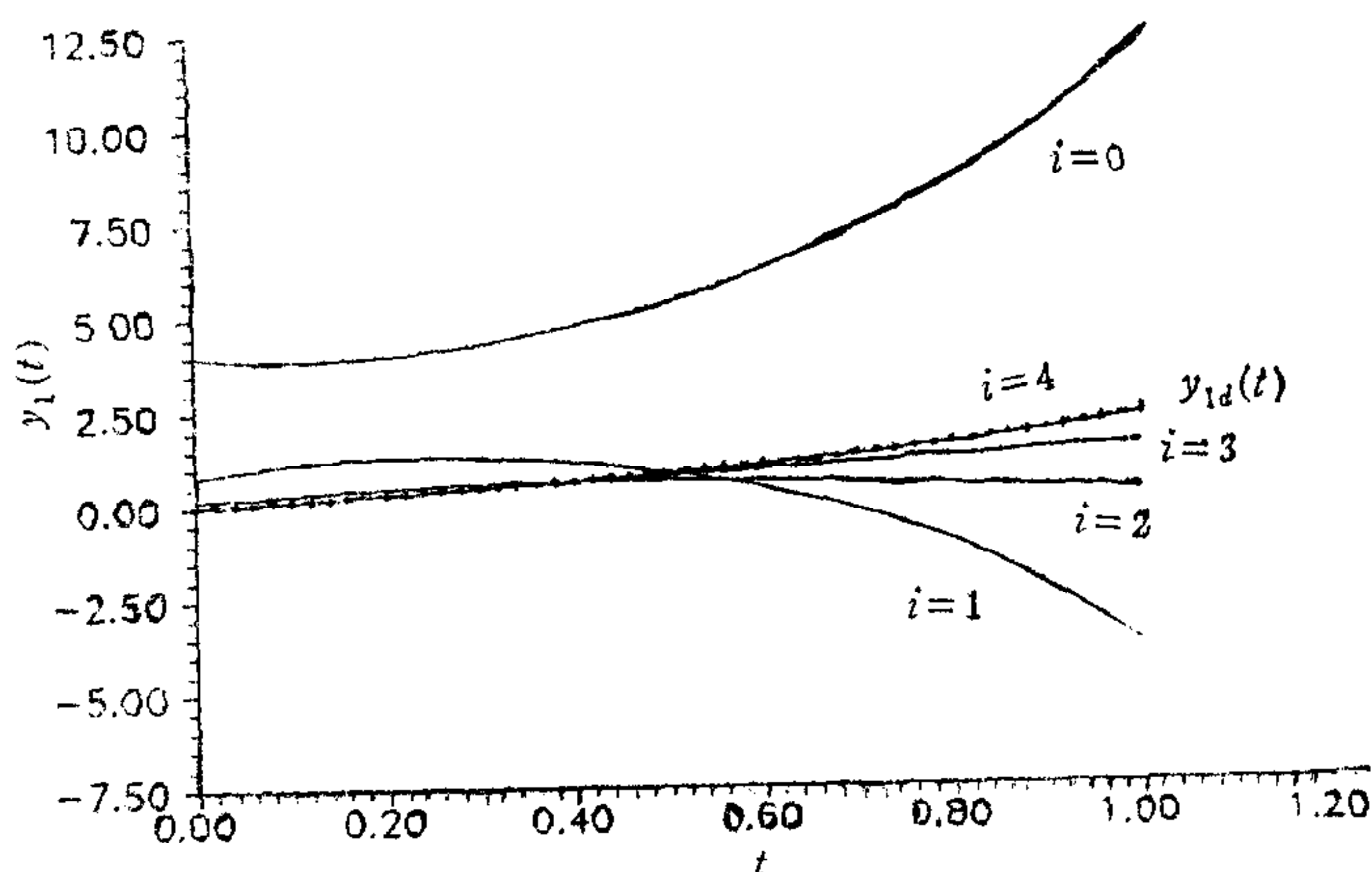


图 1 学习控制系统输出的第一分量  $y_1(t)$

设系统(4.1)学习开始时初始输入  $\begin{bmatrix} u_{1(o)}(t) \\ u_{2(o)}(t) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$  和初始状态  $\begin{bmatrix} x_{1(o)}(0) \\ x_{2(o)}(0) \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ , 经过一段时间地学习, 系统的输出曲线如图 1, 2 所示.

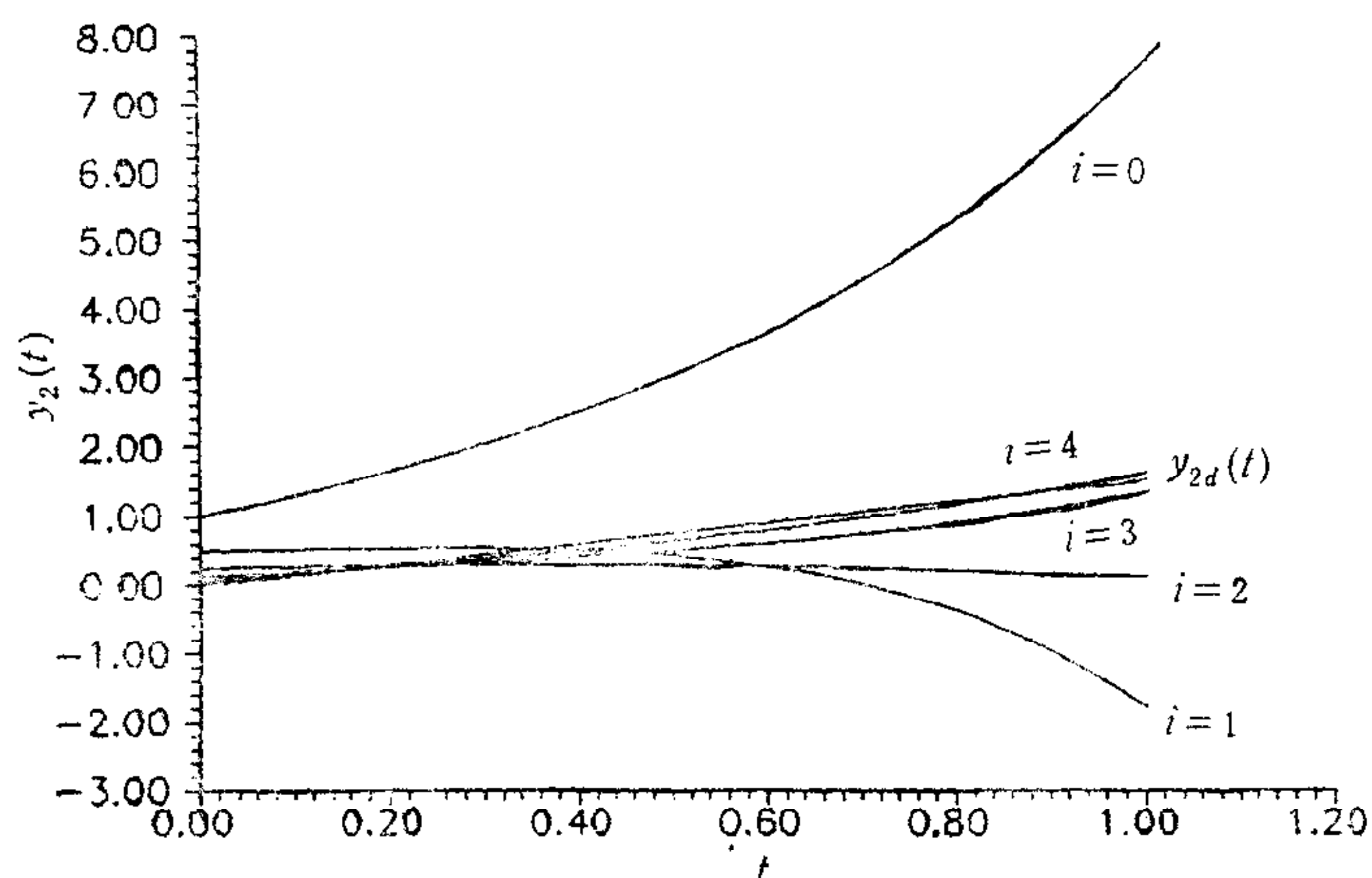


图 2 学习控制系统输出的第二分量  $y_2(t)$

上面的仿真结果表明: 开始学习时系统的初始状态  $\begin{bmatrix} x_{1(o)} \\ x_{2(o)} \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$  和期望跟踪轨迹的初始状态  $\begin{bmatrix} x_{1d}(0) \\ x_{2d}(0) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$  不等, 通过对系统的输入和初始状态的学习, 经过四次学习之后可以达到跟踪期望轨迹的目的.

## 5 结语

本文对线性时不变系统和非线性系统提出了具有初始状态学习的学习控制方案, 因而在开始学习时对系统的初始状态无要求, 即不需要假设在每次循环过程中系统的初始状态和期望跟踪轨迹的初始状态都相等. 所以应用本文的学习控制方案可以解决任意初始状态下学习控制系统的跟踪问题.

## 参 考 文 献

- [1] Arimoto S, Kawamura S and Miyazaki F. Bettering operation of dynamics systems by learning: a new control theory for servomechanism and mechatronics systems, Pro. 23rd IEEE Conf. on Decision and Control, Las Vegas, NV., 1984, 1064—1069.
- [2] Arimoto S, Kawamura S and Miyazaki F. Learning control theory for dynamical systems, Pro. 24th IEEE Conf. on Decision and Control, Fort Lauderdale, FL, 1985,, 1375—1380.
- [3] Bondi P, Casalino G and Gambardella L. On the iterative learning control theory for robotics and manipulators, *IEEE Journal of Robotics and Automation*, 1988, (1):13—22.
- [4] Hauser J E. Learning control for a class of nonlinear systems, Pro. 26th IEEE Conf. on Decision and Control, Los Angeles, CA, 1987, 859—860.
- [5] Miller III WT. Sensor-Based control of Robotic manipulators using a general learning algorithm. *IEEE J. Robotics Autom.*, 1987 RA-3, 175—163.

- [6] Ren X M, Gao W B. On the initial conditions in leaning control, Int. IEEE Sym. on industry electronic, Xi'an China, 1992, 182—185.

## LEARNING CONTROL WITH AN ARBITRARY INITIAL STATE

REN XUEMEI GAO WEIBING

*(The Seventh Research Division Beijing University of Aeronautics and Astronautics Beijing, 100083)*

### ABSTRACT

In this paper a new learning control scheme for both linear time invariant systems and a class of nonlinear systems is studied. With input and initial state learning, we can neglect the limitation about the initial state of the dynamical system at the beginning of learning process. Finally, the effectiveness of the proposed learning algorithm is illustrated by simulation

**Key words:** linear systems; nonlinear systems; learning control; initial state.