第 12 卷第 2 期
1995 年 4 月

控制理论与应用
CONTROL THEORY AND APPLICATIONS

Vol. 12, No. 2
Apr., 1995

基于运动速度的机器人学习控制*

刘德满　刘宗富
（东北大学自动控制系·沈阳，110006）
尹朝万　吴成东
（中国科学院沈阳自动化研究所·沈阳，110015）（沈阳建筑工程学院自动化系，110015）

摘要：为了使机器人跟踪给定的期望轨线，提出了一种新的基于机器人运动重复性的学习控制法. 在这种方法中机器人通过重复试验得到期望运动，这种控制法的优点：一是对于在期望运动附近非线性机器人动力学的近似表达式的线性时变机械系统产生期望运动的输入力矩可不由估计机器人动力学的物理参数形成；二是可以适当的选择位置、速度和加速度反馈增益矩阵，从而加快误差收敛速度；三是加入了加速度反馈，减少了重复试验的次数. 这是因为在每次试验的初始时刻不存在位置和速度误差，但存在加速度误差. 另外，这种控制法的有效性通过 PUMA562 机器人的前三个关节的计算机仿真结果得到验证.

关键词：学习；试验；重复；线性化；仿真

1 引　言

尽管由机器人跟踪的期望轨线被确定描述，但精确的实现它很不容易. 主要原因是估计机械关节之间的干扰及十分影响机器人运动的不可测扰动十分困难. 实时得到这样的干扰和扰动的定量估计数据实际上不可能. 但是，似乎整个机器人动力学包括那些未知的干扰和扰动当机器人重复相同运动时可以再现. 因此，期望运动可通过很好地使用机器人运动的再现得到改进. 以前，有些学者注意到了这一点且提出了基于重复真实机器人运动的新的控制方法[1~3]. 在这种控制方法中机器人可通过试验运动的充分重复最终得到给定期望运动，象人通过重复训练一样学习期望运动. 为了在实际中实现，在当前试验中加到机器人的输入力矩仅由在前一试验中和在期望运动中的真实机器人数据简单修改. 如果对于修正律和期望运动的某些条件得到满足，那么由重复试验运动，机器人收敛到期望运动. 这种控制法的第一个优点是尽管机器人动力学的线性化时变机械系统的物理参数未知，但学习控制的收敛条件很容易得到满足；第二个优点是可以通过适当的选择位置、速度和加速度反馈增益矩阵，以加快误差收敛速度；第三个突出优点是使用加速度反馈减少了试验运动次数，这是因为在每次试验运动的初始时刻不存在位置和速度误差，但存在加速度误差.

2 机器人机械手动力学的线性化

一个由 n 个驱动器驱动的 n 关节机器人机械手的每个关节由移动或转动关节组成.

* 中国科学院机器人学开放研究实验室资助课题.
本文于 1993 年 7 月 3 日收到. 1994 年 5 月 26 日收到修改稿.

众所周知机器人的非线性动力学可表示为[4~6]

$$I(\theta)\ddot{\theta} + f(\theta,\dot{\theta}) + g(\theta) = \tau. \tag{1}$$

其中 $\theta(t) \in \mathrm{R}^n$ 表示关节角坐标，$\tau \in \mathrm{R}^n$ 表示广义力矢量. $I(\theta) \in \mathrm{R}^{n\times n}$ 表示正定、对称的惯量矩阵，$g(\theta) \in \mathrm{R}^n$ 表示重力力矩矢量. $f(\theta,\dot{\theta})$ 由离心力和哥氏力及其它由摩擦力和扰动引起的非线性元素组成.

在计算力矩控制法中，相应于期望轨线的理想关节力矩 $\tau_d(t)$ 由把期望关节角位置 $\theta_d(t)$，关节角速度 $\dot{\theta}_d(t)$，和关节角加速度 $\ddot{\theta}_d(t)$ 代入(1) 中得到

$$I(\theta_d)\ddot{\theta}_d + f(\theta_d,\dot{\theta}_d) + g(\theta_d) = \tau_d(t). \tag{2}$$

但由于外界干扰和系统参数的变化，仅使用前馈力矩控制常常会导致不稳定的响应. 因此，本文加入偏差控制部分. 对于充分小的偏差 $\delta\theta$，所需的校正力矩 $\delta\tau$ 可表示为[7]

$$M(t)\delta\ddot{\theta}(t) + E(t)\delta\dot{\theta}(t) + F(t)\delta\theta(t) = \delta\tau(t). \tag{3}$$

其中 $\delta\theta(t) = \theta(t) - \theta_d(t)$，$\delta\tau(t) = \tau(t) - \tau_d(t)$，和

$$M(t) = I(\theta_d(t)),$$

$$E(t) = \left[\frac{\partial f(\theta,\dot{\theta})}{\partial\dot{\theta}}\right]_{(\theta_d,\dot{\theta}_d)},$$

$$F(t) = \left[\frac{\partial I(\theta)}{\partial\theta}\right]_{\theta_d}\ddot{\theta}_d + \left[\frac{\partial f(\theta,\dot{\theta})}{\partial\theta}\right]_{(\theta_d,\dot{\theta}_d)} + \left[\frac{\partial g(\theta)}{\partial\theta}\right]_{\theta_d}.$$

考虑由(3) 描述的线性时变系统的控制律. 为此，提出一种新的综合反馈控制和学习控制的控制法

$$U(t) = K_P(\theta_d - \theta) + K_v(\dot{\theta}_d - \dot{\theta}) + K_a(\ddot{\theta}_d - \ddot{\theta}) + u(t). \tag{4}$$

其中 K_P，K_v 和 K_a 分别为关节角位置、速度和加速度反馈增益矩阵，它们都是对角的、正定的常数矩阵，$u(t)$ 为用于学习的控制部分. 如果让 $x(t) = \theta(t) - \theta_d(t)$，$U(t) = \delta\tau(t)$. 那么，综合(4) 和(3)，可得

$$(M(t) + K_a)\ddot{x}(t) + (E(t) + K_v)\dot{x}(t) + (F(t) + K_P)x(t) = u(t). \tag{5}$$

因为 $M(t)$ 是对称、正定的惯量矩阵，K_a 是对称、正定的常数矩阵. 那么，$M(t) + K_a$ 也是对称的正定矩阵.

3 线性时变机械系统的学习控制

为了解释提出的学习控制法的实质，考虑一个线性时变机械系统

$$\begin{aligned} R(t)\ddot{x}(t) + Q(t)\dot{x}(t) + P(t)x(t) &= u(t), \\ y(t) &= \dot{x}(t). \end{aligned} \tag{6}$$

其中 $u(t) \in \mathrm{R}^n$ 和 $y(t) \in \mathrm{R}^n$ 分别是输入和输出，$x(t) \in \mathrm{R}^n$ 是位置矢量. 我们不必知道系数矩阵 $R(t)$，$Q(t)$ 和 $P(t)$ 的精确值，仅需假定矩阵 $R(t)$ 对于所有 $t \in [0,T]$(T 为终端时间) 是对称和正定的. 另外，假定这些矩阵的每个元素连续可微. 这里，为了解释由于重复试验速度收敛到期望值，设定 $y = \dot{x}$. 假定对于这个系统在 $[0,T]$ 上连续可微的期望输出 $y_d(t)$ 被给定，相应于 y_d 的期望位置由 x_d 表示. 因为系数矩阵 $R(t)$，$Q(t)$ 和 $P(t)$ 未知，产生期望输出的输入不可能通过式(6) 的计算得到. 那么我们引进一种学习方法以实现相应于机器

人期望运动的期望输出. 这可通过下列方法进行. 在首次试验中,在$[0,T]$上连续的适当的输入$u_1(t)$被给定到系统. 由输入$u_1(t)$控制的系统受下列微分方程的约束

$$\begin{aligned} &R(t)\ddot{x}_1(t) + Q(t)\dot{x}_1(t) + P(t)x_1(t) = u_1(t), \\ &y_1(t) = \dot{x}_1(t). \end{aligned} \tag{7}$$

设定初始值为

$$\dot{x}_1(0) = y_d(0) = \dot{x}_d(0), \quad x_1(0) = x_d(0). \tag{8}$$

实际上,为了速度的初始条件容易满足,通常选择初始速度为零的期望运动. 因为通常实际输出不与期望值重合,所以在第二次试验的输入$u_2(t)$由$u_1(t)$和误差$e_1(t)$形成,$e_1(t)$是期望输出$y_d(t)$与系统在第一次试验的输出$y_1(t)$之间的误差,$u_2(t)$以这种形式给出

$$u_2(t) = u_1(t) + \Lambda e_1(t), \quad e_1(t) = y_d(t) - y_1(t). \tag{9}$$

其中Λ是一个正定常数矩阵. 在第二次试验中,输入$u_2(t)$以相同的初始条件输入给系统. 在第二次试验后相同的操作被重复. 这个学习控制算法通常表示为

$$\begin{aligned} &R(t)\ddot{x}_k(t) + Q(t)\dot{x}_k(t) + P(t)x_k(t) = u_k(t), \\ &\dot{x}_k(0) = y_d(0) = \dot{x}_d(0), \quad x_k(0) = x_d(0), \\ &u_{k+1}(t) = u_k(t) + \Lambda e_k(t), \\ &e_k(t) = y_d(t) - y_k(t). \end{aligned} \tag{10}$$

学习控制如图1所示,这里问题来自于通过这种重复操作误差能否减小. 为了回答这个问题,进行下列推导定理证明.

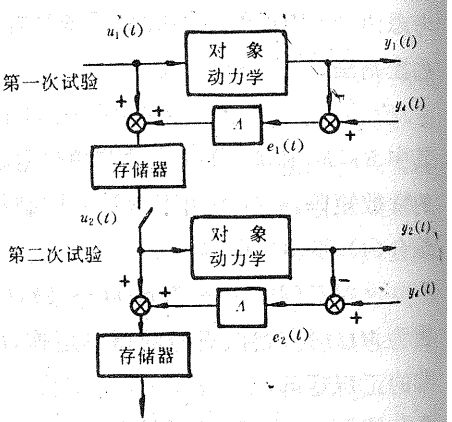

图1 学习控制

从(10),下列方程在$(k+1)$次试验成立

$$\begin{aligned} &R(t)\ddot{x}_{k+1}(t) + Q(t)\dot{x}_{k+1}(t) + P(t)x_{k+1}(t) \\ &\quad = u_{k+1}(t) \\ &\quad = u_k(t) + \Lambda e_k(t). \end{aligned} \tag{11}$$

从(11)减(10)得

$$R(t)\ddot{d}_k(t) + Q(t)\dot{d}_k(t) + P(t)d_k(t) = \Lambda e_k(t). \tag{12}$$

其中$d_k(t) = x_{k+1}(t) - x_k(t)$. 因$u_1(t)$的每个分量连续,$y_d(t)$的每个分量连续可微,而且,矩阵$R(t)$,$Q(t)$和$P(t)$的每个元素连续可微,即$u_1 \in C[0,T]$,$y_d(t) \in C^1[0,T]$,$R(t),Q(t),P(t) \in C^1[0,T]$,得到

$$x_1(t) \in C^2[0,T], \tag{13}$$

$$e_1(t) = y_d(t) - y_1(t) = y_d(t) - \dot{x}_1(t) \in C^1[0,T]. \tag{14}$$

于是,从(12)有

$$d_1(t) \in C^3[0,T], \tag{15}$$

这意味着

$$x_2(t) = d_1(t) + x_1(t) \in C^2[0,T]. \tag{16}$$

于是得到

$$e_2(t) = y_d(t) - y_2(t) = y_d(t) - \dot{x}_2(t) \in C^1[0,T].$$

在第二次试验后,同样的操作被重复,有

$$e_k(t) \in C^1[0,T], \quad d_k(t) \in C^3[0,T]. \tag{17}$$

对于任何试验次数 k，将(12)微分得到关系

$$\frac{\mathrm{d}}{\mathrm{d}t}[R(t)\ddot{d}_k(t) + Q(t)\dot{d}(t) + P(t)d_k(t)] = \Lambda\dot{e}_k. \tag{18}$$

对于任何试验次数 k，叙述下列定理.

定理 1　如果 $u_1(t)$ 的每个分量连续，$y_d(t)$ 的每个分量连续可微，那么由(10)给出的学习控制方法中的误差 $e_k(t)$ 以下列形式递减

$$J_k \geqslant J_{k+1}. \tag{19}$$

其中 J_k 由下列方程定义

$$J_k = \int_0^T e^{-\rho t}[e_k^{\mathrm{T}}(t)\Lambda e_k(t) + \lambda\dot{e}_k^{\mathrm{T}}(t)\Lambda\dot{e}_k(t)]\mathrm{d}t. \tag{20}$$

选择适当的正常数标量 ρ 和 λ，证明见附录 A.

上述定理保证随着试验次数的增加，误差不增加. 但是，实际上当试验次数趋于无穷时期望误差收敛到零，这可由从定理 1 演绎出的下列定理 2 证明.

定理 2　在定理 1 的相同条件下，误差以下列收敛

$$e_k(t) \to 0, \quad k \to \infty, \tag{21}$$

即

$$y_k(t) \to y_d(t) = \dot{x}_d(t). \tag{22}$$

对于任何固定时间 $t \in [0,T]$，证明见附录 B.

注意到初始位置设定为期望值，上述定理意味着对于任何固定时间 $t \in [0,T]$，$x_k(t)$ 收敛到期望位置 $x_d(t)$.

4　学习控制在机器人机械手中的应用

我们将所提出的学习控制法应用到具有由 n 个驱动器驱动的 n 关节的机器人机械手中去. 考虑前面所述理论是否可用于由(5)表示的线性时变系统，式(5)可重写为

$$R(t)\ddot{x}(t) + Q(t)\dot{x}(t) + P(t)x(t) = u(t). \tag{23}$$

其中

$$R(t) = M(t) + K_a; \quad Q(t) = E(t) + K_v; \quad P(t) = F(t) + K_p.$$

在这种情况下，因为 $x(t)$ 代表偏差，期望输出 $y_d(t)$ 必须设定为

$$y_d(t) = \dot{x}_d(t) = \dot{\theta}_d(t) - \dot{\theta}_d(t) = 0. \tag{24}$$

对于所有 $t \in [0,T]$. 于是，$y_d(t)$ 明显的连续可微. 其次，在第一次试验的输入 $u_1(t)$ 给定为

$$u_1(t) = 0. \tag{25}$$

对于所有 $t \in [0,T]$. 回忆一下(5)中的矩阵 $M(t)$，$E(t)$ 和 $F(t)$ 的定义，知道如果设定 $\theta_d(t) \in C^3[0,T]$ 的每个分量，那么 $u_1(t)$ 连续，$M(t)$，$E(t)$ 和 $F(t)$ 连续可微. 另外，设定初始条件为

$$\begin{aligned} \dot{x}_k(0) &= y_k(0) = \dot{\theta}_k(0) - \dot{\theta}_d(0) = 0. \\ x_k(0) &= \theta_k(0) - \theta_d(0) = 0. \end{aligned} \tag{26}$$

于是，在定理 1 和定理 2 中所要求的对于期望输出 $y_d(t)$，首次试验输入 $u_1(t)$，初始值 $x_k(0)$

和 $\dot{x}_k(0)$ 得到了满足. 另外，因 $M(t)$ 是对称、正定矩阵，K_a 是对称的正常数矩阵，那么 $(M(t)+K_a)$ 也是对称、正定矩阵. 矩阵 $M(t)$，$E(t)$ 和 $F(t)$ 的每个元素连续可微，这样定理 1 和 2 的所有条件得到了满足.

5 计算机仿真

在理论上已经表明通过重复试验机器人运动接近给定的期望运动. 但是，必须注意到当(20)中标量 ρ 被设定十分大时，在时间周期$[0,T]$的终了时间 t 与初始时间相比，误差 $e(t)$ 收敛到零较慢，需要许多试验以一致减小误差. 从实际的观点看，表明对于机器人机械手用一个十分小的 ρ 值保证误差收敛十分重要. 如果速度反馈的增益矩阵 K_v 的元素与其它矩阵 $M(t)$，$E(t)$，$F(t)$，和 Λ 的元素对于所有时间 $t\in[0,T]$ 相比被设定为足够大，从理论的观点看这很容易表示. 但反馈增益矩阵 K_v 的元素对于实际使用的机器人是否可以设定为足够大还不清楚. 为了证实实际有效性，所提出的方法被应用到 PUMA562 机器人机械手，机器人的物理参数给定为

采样时间： $t=6.0\text{ms}$.

反馈增益： $K_p=\text{diag}\ (64,64,64)\ \text{N/rad}$， $K_v=\text{diag}\ (32,32,32)\ \text{Ns/rad}$，

$K_a=\text{diag}(16,16,16)\ \text{Ns}^2/\text{rad}$.

初始条件： $(x_d(0),y_d(0),z_d(0))=(0.45,0.25,0.04)\text{m}$.

终止条件： $(x_d(T),y_d(T),z_d(T))=(0.45,0.25,0.04)\text{m}$.

终止时间： $T=2\text{s}$.

期望运动： $x_d(t)=0.25+0.2\cos(\pi t)$， $y_d(t)=0.25+0.2\sin(\pi t)$，

$z_d=0.04\text{m}$.

学习增益： $\Lambda=\text{diag}(32,32,32)\text{Ns/rad}$.

从图 2 至图 5 可以看出：经过 4 次试验跟踪误差趋近于零. 可见，本文提出的学习控制法具有快速收敛的优点.

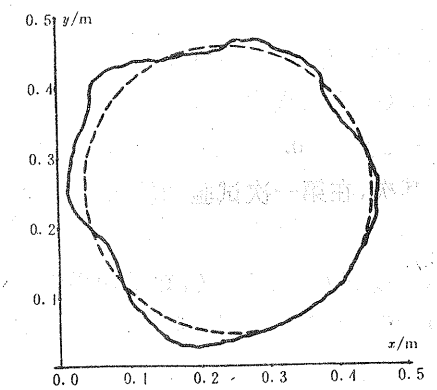

图 2 第一次试验结果

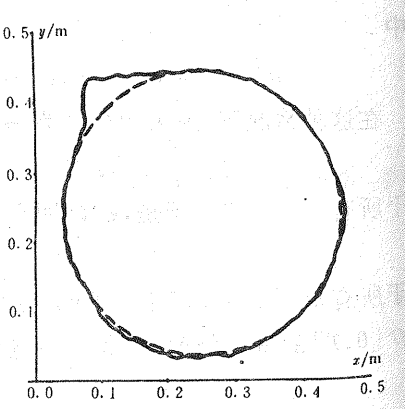

图 3 第二次试验结果

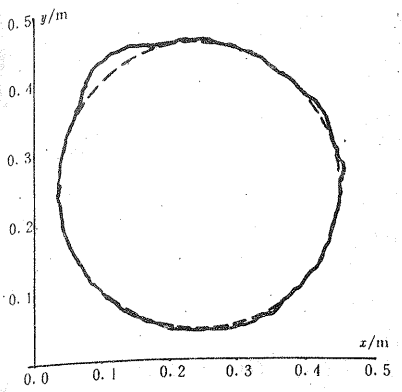

图 4 第三次试验结果

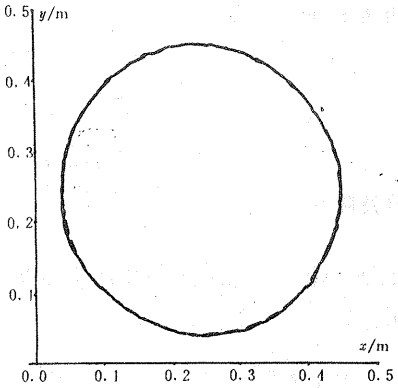

图 5 第四次试验结果

参 考 文 献

[1] Arimoto, S., Kawamura, S. and Miyazaki, F.. Bettering Operation of Robots by Learning. J. Robotic Sys., 1984, 1: 123—128

[2] Arimoto, S., Kawamura, S. and Miyazaki, F.. Bettering Operation of Dynamic Systems by Learning: A New Control Theory for Servomechanics or Mechatronics Systems. in Proc. 23rd IEEE Conf. Decision and Control, Las Vegas, NV, Dec. 1984

[3] Arimoto, S., Kawamura, S. and Miyazaki, F.. Can Mechanical Robots Learn by Themselves? in Robotics Research, The 2nd International Symposium. Cambridge, MA: MTT Press, 1985, 127—132

[4] Mclnroy, J. E. and Saridis, G. N.. Acceleration and Torque Feedback for Robotic Control: Experimental Results. J. Robotic Sys., 1990, 6: 813—832

[5] Arimoto, S. and Miyazaki, F.. Stabiiity and Robustness of PID Feedback Control for Robot Manipulator of Sensory Capabilities. in Robotics Research, The lst International Symposium. Cambridge, MA: MIT Press, 1983

[6] Takegaki, M and Arimoto, S.. A New Feedback Method for Dynamic Control of Manipulators. Trans. ASME, J. Dynamic Sys. Meas. Contr., 1981, 103: 119—125

[7] Lee, C. S. G., and Chung, M. J.. An Adaptive Control Strategy for Mechanical Manipulators. IEEE Trans. Automat. Contr., 1984, AC-29:837—840

[8] Kawamura, S., and Miyazaki, F., and Arimoto, S.. Hybrid Position/Force Control of Robot Manipulators Based on Learning Method. in Proc. ICAR 85, Tokyo, Japan, Sept. 1985

[9] Kawamura, S., Miyazaki, F., and Arimoto, S.. Application of Learning Methods for Dynamic Controls of Robot Manipulator. in Proc. 24th IEEE Conf. Decision and Control, Fort Lauderdale, FL., Dec. 1985

附录 A

定理 1 的证明

由 $d_k(t)$ 的定义得

$$\begin{aligned}\dot{d}_k(t) &= \dot{x}_{k+1}(T) - \dot{x}_k(t)\\ &= y_{k+1}(t) - y_k(t)\\ &= y_d(t) - y_k(t) - (y_d(t) - y_{k+1}(t))\\ &= e_k(t) - e_{k+1}(t),\end{aligned} \tag{A1}$$

这同样意味着

$$\ddot{d}_k(t) = \dot{e}_k(t) - \dot{e}_{k+1}(t). \tag{A2}$$

将(A1)和(A2)代入由(20)定义的 J_{k+1} 得到

$$\begin{aligned}J_{k+1} &= \int_0^T e^{-\rho t}[e_{k+1}^{\mathrm{T}}(t)\Lambda e_{k+1}(t) + \lambda\dot{e}_{k+1}^{\mathrm{T}}(t)\Lambda\dot{e}_{k+1}(t)]\mathrm{d}t\\ &= \int_0^T e^{-\rho t}[(e_k(t) - \dot{d}(t))^{\mathrm{T}}\Lambda(e_k(t) - \dot{d}(t)) + \lambda(\dot{e}_k(t) - \ddot{d}_k(t))^{\mathrm{T}}\Lambda(\dot{e}(t) - \ddot{d}_k(t))]\mathrm{d}t\\ &= J_k + \int_0^T e^{-\rho t}[\dot{d}_k^{\mathrm{T}}(t)\Lambda\dot{d}_k(t) + \lambda\ddot{d}_k^{\mathrm{T}}(t)\Lambda\ddot{d}_k(t)]\mathrm{d}t - 2\int_0^T e^{-\rho t}[\dot{d}_k^{\mathrm{T}}(t)\Lambda e_k(t) + \lambda\ddot{d}_k^{\mathrm{T}}(t)\Lambda\dot{e}_k(t)]\mathrm{d}t.\end{aligned} \tag{A3}$$

此外,将(12)和(18)代入(A3)得

$$\begin{aligned}J_k - J_{k+1} &= 2\int_0^T e^{-\rho t}\dot{d}_k^{\mathrm{T}}(t)[R(t)\ddot{d}_k(t) + Q(t)\dot{d}_k(t) + P(t)d_k(t)]\mathrm{d}t\\ &\quad + 2\lambda\int_0^T e^{-\rho t}\ddot{d}_k^{\mathrm{T}}(t)\frac{\mathrm{d}}{\mathrm{d}t}[R(t)\ddot{d}_k(t) + Q(t)\dot{d}_k(t) + P(t)d_k(t)]\mathrm{d}t\\ &\quad - \int_0^T e^{-\rho t}[\dot{d}_k^{\mathrm{T}}(t)\Lambda\dot{\mathrm{d}}(t) + \lambda\ddot{d}_k^{\mathrm{T}}(t)\Lambda\ddot{d}(t)]\mathrm{d}t.\end{aligned} \tag{A4}$$

因为初始条件给定为

$$\begin{aligned}d_k(0) &= x_{k+1}(0) - x_k(0) = x_d(0) - x_d(0) = 0,\\ \dot{d}_k(0) &= \dot{x}_{k+1}(0) - \dot{x}_k(0) = \dot{x}_d(0) - \dot{x}_d(0) = 0,\\ \ddot{d}_k(0) &= \ddot{x}_{k+1}(0) - \ddot{x}_k(0) = \ddot{x}_d(0) - \ddot{x}_d(0) = 0\end{aligned} \tag{A5}$$

和

$$R(t) = M(t) + K_a,\quad Q(t) = E(t) + K_v,\quad P(t) = F(t) + K_p. \tag{A6}$$

$R(t)$ 是正定、对称矩阵,(A4)可通过部分积分写为

$$\begin{aligned}J_k - J_{k+1} &= e^{-\rho t}[\ddot{d}_k^{\mathrm{T}}(T)\lambda(M(T) + K_a)\ddot{d}_k(T) + \dot{d}_k^{\mathrm{T}}(T)(M(T) + K_a + \lambda K_p)\dot{d}_k(T) + d_k^{\mathrm{T}}K_p d_k]\\ &\quad + \lambda\int_0^T e^{-\rho t}\ddot{d}_k^{\mathrm{T}}(t)[\rho(M(t) + K_a) + 2K_v + 2E(t) + \dot{M}(t) - \Lambda]\ddot{d}_k(t)\mathrm{d}t\\ &\quad + \int_0^T e^{-\rho t}\dot{d}_k^{\mathrm{T}}(t)[\rho(M(t) + K_a + \lambda K_p) + 2K_v + 2E(t) - \dot{M}(t) - \Lambda]\dot{d}_k(t)\mathrm{d}t\\ &\quad + \int_0^T e^{-\rho t}d_k^{\mathrm{T}}[\rho K_p]d_k(t)\ \mathrm{d}t + 2\lambda\int_0^T e^{-\rho t}\ddot{d}_k^{\mathrm{T}}(t)(\dot{E}(t) + F(t))\dot{d}_k(t)\mathrm{d}t\\ &\quad + 2\lambda\int_0^T e^{-\rho t}\ddot{d}_k^{\mathrm{T}}\dot{E}(t)d_k(t)\mathrm{d}t + 2\int_0^T e^{-\rho t}\dot{d}_k^{\mathrm{T}}(t)F(t)d_k(t)\mathrm{d}t.\end{aligned} \tag{A7}$$

这里,注意以下关系很重要

$$\begin{aligned}&\int_0^T e^{-\rho t}[\dot{d}_k(t) + F(t)d_k(t)]^{\mathrm{T}}[\dot{d}_k(t) + F(t)d_k(t)]\mathrm{d}t \geqslant 0,\\ &\lambda\int_0^T e^{-\rho t}[\ddot{d}_k(t) + \dot{F}(t)d_k(t)]^{\mathrm{T}}[\ddot{d}_k(t) + \dot{F}(t)d_k(t)]\mathrm{d}t \geqslant 0,\\ &\lambda\int_0^T e^{-\rho t}[\ddot{d}_k(t) + (\dot{F}(t) + F(t))\dot{d}_k(t)]^{\mathrm{T}}[\ddot{d}_k(t) + (\dot{E}(t) + F(t))\dot{d}_k(t)]\mathrm{d}t \geqslant 0.\end{aligned} \tag{A8}$$

λ 和 ρ 是正标量,显然有

$$\begin{aligned}&\int_0^T e^{-\rho t}\dot{d}_k^{\mathrm{T}}(t)\dot{d}_k(t)\mathrm{d}t + \int_0^T e^{-\rho t}d_k^{\mathrm{T}}(t)F^{\mathrm{T}}(t)F(t)d_k(t)\mathrm{d}t + 2\int_0^T e^{-\rho t}\dot{d}_k^{\mathrm{T}}(t)F(t)d_k(t)\mathrm{d}t \geqslant 0,\\ &\lambda\int_0^T e^{-\rho t}\ddot{d}_k^{\mathrm{T}}(t)\ddot{d}(t)\mathrm{d}t + \lambda\int_0^T e^{-\rho t}d_k^{\mathrm{T}}(t)\dot{F}^{\mathrm{T}}(t)\dot{F}(t)d_k(t)\mathrm{d}t + 2\lambda\int_0^T e^{-\rho t}\ddot{d}_k^{\mathrm{T}}(t)\dot{F}(t)d_k(t)\mathrm{d}t \geqslant 0,\end{aligned}$$

$$\lambda\int_0^T e^{-\rho t}\ddot{d}_k^{\mathrm{T}}(t)\ddot{d}_k(t)\mathrm{d}t+\lambda\int_0^T e^{-\rho t}\dot{d}_k^{\mathrm{T}}(t)(\dot{E}(t)+F(t))^{\mathrm{T}}(\dot{E}(t)+F(t))\dot{d}_k(t)\mathrm{d}t$$
$$+2\lambda\int_0^T e^{-\rho t}\ddot{d}_k^{\mathrm{T}}(t)(\dot{E}(t)+F(t))\dot{d}_k\geqslant 0. \tag{A9}$$

将(A9)代入不等式(A7)的右边，可得

$$\begin{aligned}J_k-J_{k+1}\geqslant e^{-\rho t}&[\ddot{d}_k^{\mathrm{T}}(T)\lambda(M(T)+K_a)\ddot{d}(T)\\&+\dot{d}_k^{\mathrm{T}}(T)(M(T)+K_a+\lambda K_p)\dot{d}_k(T)+d_k^{\mathrm{T}}(T)K_pd_k(T)]\\&+\int_0^T e^{-\rho t}\ddot{d}_k^{\mathrm{T}}(t)A(t)\ddot{d}_k(t)\mathrm{d}t+\int_0^T e^{-\rho t}\dot{d}_k^{\mathrm{T}}(t)B(t)\dot{d}_k(t)\mathrm{d}t+\int_0^T e^{-\rho t}d_k^{\mathrm{T}}(t)C(t)d_k(t)\mathrm{d}t.\end{aligned} \tag{A10}$$

其中

$$\begin{aligned}A(t)&=\lambda[\rho(M(t)+K_a)+2K_v+2E(t)+\dot{M}(t)-\Lambda-2I],\\B(t)&=\rho(M(t)+K_a+\lambda K_p)+2K_v+2E(t)-\dot{M}(t)-\Lambda-I\\&\quad-\lambda(\dot{E}(t)+F(t))^{\mathrm{T}}(\dot{E}(t)+F(t)),\\C(t)&=\rho K_pI-\lambda\dot{F}^{\mathrm{T}}(t)\dot{F}(t)-F^{\mathrm{T}}(t)F(t).\end{aligned} \tag{A11}$$

其中 I 是单位矩阵. 如果

$$A(t)>0,\quad B(t)>0,\quad C(t)>0, \tag{A12}$$

这意味着式(A10)的值是非负的. 因此可作出结论

$$J_k\geqslant J_{k+1}, \tag{A13}$$

那么误差的收敛得到保证. 如果反馈增益矩阵 K_a，K_v 和 K_p 的每个对角项设定得比系数矩阵和它们的微分的其它元素大得多，那么可见(A11)对于十分小的 ρ 值得到满足. 因此我们可通过选择适当的 K_a，K_v 和 K_p，以便对于比较小的 ρ 误差的收敛得到保证.

附录 B

定理 2 的证明

因为定理中所有关系成立尽管我们设定 $t\in[0,T]$ 代替 T，于是我们得到下列不等式代替不等式(A10)

$$\begin{aligned}J_k(t)-J_{k+1}\geqslant e^{-\rho t}&[\ddot{d}_k^{\mathrm{T}}(t)\lambda(M(t)+K_a)\ddot{d}_k(t)\\&+\dot{d}_k^{\mathrm{T}}(t)(M(t)+K_a+\lambda K_p)\dot{d}_k(t)+d_k^{\mathrm{T}}(t)K_pd_k(t)]\\&+\int_0^t e^{-\rho t}\ddot{d}_k^{\mathrm{T}}(\tau)A(\tau)\ddot{d}_k(\tau)\mathrm{d}\tau+\int_0^t e^{-\rho t}\dot{d}_k^{\mathrm{T}}(\tau)B(\tau)\dot{d}_k(\tau)\mathrm{d}\tau\\&+\int_0^t e^{-\rho t}d_k^{\mathrm{T}}(\tau)C(\tau)d_k(\tau)\mathrm{d}\tau\geqslant 0.\end{aligned} \tag{B1}$$

对于任何固定时间 $t\in[0,T]$. 从(B1)，有

$$\begin{aligned}&J_k(t)-J_{k+1}(t)\geqslant e^{-\rho t}\ddot{d}_k^{\mathrm{T}}(t)\lambda(M(t)+K_a)\ddot{d}_k(t)\geqslant 0,\\&J_k(t)-J_{k+1}(t)\geqslant e^{-\rho t}\dot{d}_k^{\mathrm{T}}(t)(M(t)+K_a+\lambda K_p)\dot{d}_k(t)\geqslant 0,\\&J_k(t)-J_{k+1}(t)\geqslant e^{-\rho t}d_k^{\mathrm{T}}(t)K_pd_k(t)\geqslant 0\end{aligned} \tag{B2}$$

对于任何固定时间 $t\in[0,T]$，因为式(B1)的右边的每项是正定的. 系数矩阵 $R(t)$ 是正定的，J_k 对于任何固定时间收敛到某个值，于是可作出结论：

$$\ddot{d}_k(t)\to 0,\quad \dot{d}_k(t)\to 0,\quad d_k(t)\to 0,\quad \text{当 } k\to\infty. \tag{B3}$$

对于任何固定时间 $t\in[0,T]$. 从(12)，这意味着

$$e_k(t)\to\infty,\text{当 } k\to\infty. \tag{B4}$$

对于任何固定时间 $t\in[0,T]$.

Robot Learn Control Based-On Motion Rate

LIU Deman and LIU Zongfu
(Department of Automatic Control, Northeastern University • Shenyang, 110006, PRC)
YIN Chaowan
(Robotics Laboratory, Shenyang Institute of Automation • Shenyang, 110015, PRC)
WU Chendong
(Department of Automatic Control, Shenyang Civil Architecture Engineering Institute • Shenyang, 110015, PRC)

Abstract: To make a robot track a given desired motion trajectory, a new learning control scheme is proposed which is based on the repeatability of robot motion. In this scheme the robot obtains a desired motion by repeating trials. This control method has three advantages: First, this control scheme is that the input torque that generates the desired motion can be formed without estimating the physical parameters of a linear time-varying mechanical system, which is an approximate representation of nonlinear robot dynamics in the vicinity of the desired motion; Second, the rate of error convergence may be guaranteed rapidly by setting appropriate angular position, velocity and acceleration feedback gain Matrices; Third, in this control method, acceleration feedback is added to decrease the number of repeated trials. This is duo to the fact that no position and velocity error exist at the beginning of every trial movement. Moreover, effectiveness of this control scheme is demonstrated through computer simulation in which a PUMA 562 manipulator with first three degrees of freedom is used.

Key words: learning; trial; repeated; linearized; computer simulation

本文作者简介

刘德满 1963年生，副教授. 1984年在武汉钢铁学院电气化系获工学学士学位. 后分别于1986年和1991年在东北大学自动控制系获工学硕士、博士学位. 主要研究工作兴趣是非线性控制，自适应控制，机器人控制. 目前研究领域是机器人智能控制.

刘宗富 1925年生. 教授. 1950年毕业于同济大学电机工程系，1950年到东北大学任教，1956年晋升为讲师，1958年获苏联副博士学位. 1978年晋升为副教授，1983年晋升为教授. 现为中国自动化学会电气自动化专业委员会委员.

尹朝万 1940年生. 研究员. 1962年毕业于华中理工大学自动控制系，1962年开始在北京中科院自动化研究所工作. 1965年到中国科学院沈阳自动化研究所工作，现任机器人学开放研究实验室副主任. 主要研究兴趣是数据库技术和智能控制等. 现在的研究领域为机器智能控制和CIMS信息集成.

吴成东 1960年生. 副教授. 1983年毕业于沈阳建筑工程学院自动控制系，并留校任教. 1985年考入清华大学自动化系攻读硕士学位，1988年获工学硕士学位. 1991年考入东北大学自动控制系攻读在职博士学位. 目前正从事机器人智能控制方面的研究.