講座

機械システムの知能化 V: 偏差ダイナミクス, モデルベース適応制御, 学習制御

To Make Mechanical systems More Intelligent V: Residual Dynamics, Learning Control, and Model-based Adaptive Control

有本 卓* 東京大学工学部

Suguru Arimoto* *Faculty of Engineering Science, University of Tokyo

1. はじめに

今までは目標姿勢を与えたときの定置制御を考えたが、今度は、理想の運動が関節座標ベクトルの時間推移 $q_d(t)$ で与えられたときの軌道追従の問題を考え、その制御入力を定めるための二つの考え方を講述する。一つはモデルベース適応制御であり、他の一つは学習制御である。前者ではマニピュレータの主要パラメータが未知でもよいとして、それらを推定しつつ、時間の経過とともにロボットの運動を表す関節座標ベクトル $q(t)$ を理想のそれ $q_d(t)$ に漸近的に追従させたい。後者では、ダイナミクスがほぼ未知であっても、何回かの繰り返し練習の後、理想の運動がほぼ実現できるようにしたい。両方とも、偏差ダイナミクスの受動性が本質的な役割を演じる。

2. 偏差ダイナミクスの受動性

ロボットマニピュレータの理想の運動が関節座標ベクトル q の時間経過 $q_d(t)$ で与えられているとしよう。ここに $q_d(t)$ は 2 回微分可能であり、 $\ddot{q}_d(t)$ は区分的連続かつ有界、また $\dot{q}_d(t)$ も有界とする。今、マニピュレータには SP-D サーボが実装してある、すなわち、

$$F = -B_1 \dot{q} - A_1 s(\Delta q) + u \quad (1)$$

が組み込まれているとする。ここに $\Delta q = q - q_d$ であり、 $s(\Delta q)$ の第 i 成分 $s_i(\Delta q_i)$ は図 III-5 に示したような飽和型の関数から成るとする。また、 u はこれから決めるべき外部ループの制御入力とする。式 (1) を式 (II-14) に代入すると、式

$$\begin{aligned} \{H_0 + H(q)\} \ddot{q} + \left\{ B + \frac{1}{2} \dot{H}(q) + S(q, \dot{q}) \right\} \dot{q} + g(q) \\ + A_1 s(\Delta q) = u \end{aligned} \quad (2)$$

を得る。ここに $B = B_0 + B_1$ である。ここで、理想の運動 $q_d(t)$ を与える入力トルクを $u_d(t)$ とすると、それは式 (2) から次のように書けねばならない。

$$\begin{aligned} \{H_0 + H(q_d)\} \ddot{q}_d + \left\{ B + \frac{1}{2} \dot{H}(q_d) + S(q_d, \dot{q}_d) \right\} \dot{q}_d \\ + g(q_d) = u_d \end{aligned} \quad (3)$$

式 (2) から (3) を引くと

$$\begin{aligned} \{H_0 + H(q)\} \Delta \ddot{q} + \left\{ B + \frac{1}{2} \dot{H}(q) + S(q, \dot{q}) \right\} \Delta \dot{q} \\ + A_1 s(\Delta q) + h(\Delta q, \Delta \dot{q}) = \Delta u \end{aligned} \quad (4)$$

となる。ここに $\Delta u = u - u_d$ であり、 h は次のように表される。

$$\begin{aligned} h(\Delta q, \Delta \dot{q}) = \{H(q) - H(q_d)\} \ddot{q}_d \\ + \left\{ \frac{1}{2} \dot{H}(q) + S(q, \dot{q}) - \frac{1}{2} \dot{H}(q_d) - S(q_d, \dot{q}_d) \right\} \dot{q}_d \\ + g(q) - g(q_d) \end{aligned} \quad (5)$$

右辺の第 2 項の $\{ \}$ の中は

$$\begin{aligned} \sum_{i=1}^n H_i(q) (\dot{q}_i - \dot{q}_{di}) + S(q, \dot{q}) - S(q, \dot{q}_d) \\ + \sum_{i=1}^n \{H_i(q) - H_i(q_d)\} \dot{q}_{di} + S(q, \dot{q}_d) - S(q_d, \dot{q}_d) \end{aligned}$$

と表され、この前半部は $\Delta \dot{q}$ について線形斉次の関係にあり、後半部は Δq のオーダーにあることが判る。ここに $H_i(q)$ は $H(q)$ の q_i による偏微分を表す。こうして、式 (5) は

$$h(\Delta q, \Delta \dot{q}) = E(q_d, \dot{q}_d) \Delta \dot{q} + r(q_d, \dot{q}_d, \ddot{q}_d; \Delta q) \quad (6)$$

と表されることが判る。ここに r は Δq について 1 次のオーダーであり、しかも $H(q)$ と $g(q)$ の各成分が三角関数から成るので Δq について有界でもある。従って明らかに

$$\Delta \dot{q}^T h(\Delta q, \Delta \dot{q}) \geq -c_0 \|\Delta \dot{q}\|^2 - c_1 \|s(\Delta q)\|^2 \quad (7)$$

原稿受付 1994 年 5 月 31 日

キーワード: Residual Robot Dynamics, Model-Based Adaptive Control, Learning Control, Repetitive Control, Passivity

〒113 東京都文京区本郷 7-3-1

*Bunkyo-ku, Tokyo

を満たす正の定数 c_0, c_1 が存在する. ここに $s(\Delta q)$ は $\|\Delta q\|$ が小さいとき $\|\Delta q\|$ のオーダーであることに注意しておく. 他方,

$$s(\Delta q)^T h(\Delta q, \Delta \dot{q}) \geq -c_2 \|\Delta \dot{q}\|^2 - c_3 \|s(\Delta q)\|^2 \quad (8)$$

となる正定数 c_2, c_3 も存在する. しかも, これらの定数は $H(q)$ の最大特異値と $q_d, \dot{q}_d, \ddot{q}_d$ にのみ依存しており, 決して高利得に取る必要はないことに注意しておく.

さて, 出力を速度の偏差と位置偏差の飽和関数との線形和

$$\Delta z = \Delta \dot{q} + \alpha s(\Delta q) \quad (9)$$

に取り, これと偏差ダイナミクス (4) との間の受動性を検討してみよう. まず, 明らかに

$$\begin{aligned} \Delta \dot{q}^T \Delta u \geq & \frac{d}{dt} \left[\frac{1}{2} \Delta \dot{q}^T \{H_0 + H(q)\} \Delta \dot{q} + \sum_{i=1}^n a_{1i} S_i(\Delta q_i) \right] \\ & + \Delta \dot{q}^T (B - c_0 I) \Delta \dot{q} - c_1 \|s(\Delta q)\|^2 \end{aligned} \quad (10)$$

である. ここに a_{1i} は A_1 の第 i 番目の対角成分を表す. また,

$$\begin{aligned} s^T(\Delta q) \left\{ -\frac{1}{2} \dot{H}(q) + S(q, \dot{q}) \right\} \Delta \dot{q} \\ \geq -c_4 \|\Delta \dot{q}\|^2 - c_5 \|s(\Delta q)\|^2 \end{aligned} \quad (11)$$

が成立するような正定数 c_4 と c_5 が存在することも容易に確かめられる. 式 (11) より

$$\begin{aligned} s^T(\Delta q) \Delta u \\ = & \frac{d}{dt} \left[s^T(\Delta q) \{H_0 + H(q)\} \Delta \dot{q} + \sum_{i=1}^n b_i S_i(\Delta q_i) \right] \\ & - \left\{ \frac{d}{dt} s(\Delta q) \right\}^T H(q) \Delta \dot{q} \\ & + s^T(\Delta q) \left\{ -\frac{1}{2} \dot{H}(q) + S(q, \dot{q}) \right\} \Delta \dot{q} \\ & + s^T(\Delta q) h(\Delta q, \Delta \dot{q}) + s^T(\Delta q) A_1 s(\Delta q) \\ \geq & \frac{d}{dt} \left[s^T(\Delta q) \{H_0 + H(q)\} \Delta \dot{q} + \sum_{i=1}^n b_i S_i(\Delta q_i) \right] \\ & - \gamma \Delta \dot{q}^T H(q) \Delta \dot{q} - (c_2 + c_4) \|\Delta \dot{q}\|^2 \\ & - (c_3 + c_5) \|s(\Delta q)\|^2 + s^T(\Delta q) A_1 s(\Delta q) \end{aligned} \quad (12)$$

となる. ここに b_i は B の第 i 番目の対角成分を表し, γ は $s_i(\theta)$ の微分値の最大絶対値の上界を取るとする. こうして, 式 (10) と式 (12) の α 倍を加え, その結果を区間 $[0, t]$ で積分することにより, 式

$$\begin{aligned} \int_0^t \Delta z^T(\tau) \Delta u(\tau) d\tau \geq & W(\Delta q(t), \Delta \dot{q}(t)) \\ & - W(\Delta q(0), \Delta \dot{q}(0)) + \int_0^t V(\Delta q(\tau), \Delta \dot{q}(\tau)) d\tau \end{aligned} \quad (13)$$

を得る. ここに

$$\begin{aligned} W(\Delta q, \Delta \dot{q}) = & \frac{1}{2} \Delta \dot{q}^T \{H_0 + H(q)\} \Delta \dot{q} \\ & + \sum_{i=1}^n (a_{1i} + \alpha b_i) S_i(\Delta q_i) + \alpha s^T(\Delta q) \{H_0 + H(q)\} \Delta \dot{q} \end{aligned} \quad (14)$$

$$V(\Delta q, \Delta \dot{q})$$

$$\begin{aligned} = & \Delta \dot{q}^T \{B - (c_0 + \alpha c_2 + \alpha c_4) I - \alpha \gamma H(q)\} \Delta \dot{q} \\ & + s^T(\Delta q) \{\alpha A_1 - \alpha(c_3 + c_5) I - c_1 I\} s(\Delta q) \end{aligned} \quad (15)$$

と置いた.

さて, 偏差ダイナミクス (4) の入出力対 $\{\Delta z, \Delta u\}$ について受動性が成立するためには, 式 (13) より, W と V がともに Δq と $\Delta \dot{q}$ について正定であれば十分である. ここで, 式 (1) のサーボのゲイン A_1 と B_1 を普通を選ぶとき, W と V が正定になり得るかどうか問題となる. 明らかに α は少し小さいくらいが望ましいが, 具体的には III の付録 A で述べた方法と同じようにして, W が正定になる条件は簡単に導ける. 実際, 今, S_i とその微分 s_i との間に

$$S_i(\theta) \geq \gamma_i s_i^2(\theta) \quad (16)$$

の関係があるとしよう ($S_i(\theta) = 1 - \cos(\theta)$ のとき, $\gamma_i = 1/2$ に取れる). そのとき,

$$\begin{aligned} & \frac{1}{4} \Delta \dot{q}^T \{H_0 + H(q)\} \Delta \dot{q} + \alpha s^T(\Delta q) \{H_0 + H(q)\} \Delta \dot{q} \\ & + \sum_{i=1}^n \alpha b_i S_i(\Delta q_i) \\ \geq & \frac{1}{4} \{\Delta \dot{q} + 2\alpha s(\Delta q)\}^T \{H_0 + H(q)\} \{\Delta \dot{q} + 2\alpha s(\Delta q)\} \\ & + \alpha s^T(\Delta q) [\gamma_0 B - \alpha \{H_0 + H(q)\}] s(\Delta q) \end{aligned} \quad (17)$$

となる. ここに $\gamma_0 = \min\{\gamma_i\}$ とした. これより, もし

$$\gamma_0 B \geq \alpha \{H_0 + H(q)\} \quad (18)$$

となるように B_1 を取れば, 式 (17) の右辺は非負となり, 従ってこのとき W は次の不等式を満足する.

$$W(\Delta q, \Delta \dot{q}) \geq \frac{1}{4} \Delta \dot{q}^T \{H_0 + H(q)\} \Delta \dot{q} + \sum_{i=1}^n a_{1i} S_i(\Delta q_i) \quad (19)$$

つまり, W は $\Delta q, \Delta \dot{q}$ について正定になる.

実は, 次節で述べるように, 一回の練習が有限時間で終わるような繰り返し学習では, W についてのみ正定性が成り立てば収束性が言えるのである. モデルベース適応制御の実時間計算を必要としない方式や, 周期的な作業に関する連続的繰り返し制御では, W のみならず V の正定性

図1 偏差ロボットダイナミクスの受動性

も必要となる。\$V\$ が正定になるためには、\$\alpha\$ が小さいとして、\$B\$ は \$c_0 I\$ よりも大きく、かつ、\$\alpha A_1\$ は \$c_1 I\$ より大きくなければならない。つまり、\$\alpha\$ をそれほど小さく取らず、\$A_1\$ と \$B_1\$ を少し大きめに取る必要がある。\$c_0\$ と \$c_1\$ は明らかに \$H(q)\$ の最大特異値と \$\dot{q}_d\$ や \$\ddot{q}_d\$ や \$g(q)\$ の成分の最大絶対値の積のオーダーであるから、アクチュエータの駆動系に減速機構がついていれば、\$H(q)\$ は減速比の 2 乗に反比例して小さくなるので、\$H_0\$ や \$A_1\$, \$B\$ の対角要素に比して小さくなり、\$V\$ は \$\Delta q\$, \$\Delta \dot{q}\$ について正定になるであろう。

結論 1 このとき、式 (4) の偏差ダイナミクスについて、式 (9) の出力 \$\Delta z\$ と入力 \$\Delta u\$ の対は受動性を満たす (図 1 参照)。

3. 学習制御の枠組

練習を繰り返すうちに理想の運動を高精度で実現させようとするのが学習制御の考え方である。そのために次のように問題を設定してみる。

- A1) 1 回の運動は時間 \$T > 0\$ で終わる。
- A2) その時間区間 \$t \in [0, T]\$ にわたって理想の出力 (運動軌道) \$y_d(t)\$ が先験的に与えられている。
- A3) 初期化は常に一定である。すなわち、これから行う練習を第 \$k\$ 回とすると、そのときの状態 \$x_k(t) = (q_k(t), \dot{q}_k(t))\$ は運動開始の際、常に同一の状態に初期化される。

$$x_k(0) = x^0, \quad (k = 0, 1, 2, \dots)$$

- A4) ダイナミクスは繰り返し練習中是不変である。
- A5) 出力軌道 \$y_k(t)\$ は測定できる。したがって、任意の \$k\$ 回目の練習の際の偏差 \$\Delta y_k = y_k(t) - y_d(t)\$ は常に計算可能である。
- A6) 次回のアクチュエータ入力 \$u_{k+1}\$ は記憶の無い関数あるいは作用素で

$$u_{k+1}(t) = F(u_k(t), \Delta z_k(t)) \quad (20)$$

と構成できる。ここに \$\Delta z_k\$ は \$\Delta y_k\$ の関数か、\$\Delta y_k\$ を入力としたある簡単なシステムの出力とする。

- A7) ダイナミクスは逆転可能である。すなわち、出力 \$y_d(t)\$ が与えられたとき、これを実現する入力 \$u_d(t)\$ は、計算できないとしても、存在することだけは仮定できるとする。

以上の条件であるものは強すぎ、現実には満足できないものもあるが、それを緩和したときの問題は後で論じる。

人の技能の習熟過程を観察すると、それは決して目覚しい速さでは進んでいない。動物の学習はむしろ遅いこと (tediousness) に特徴がある。しかし、練習すれば習熟が確実に進む。その秘密は偏差ダイナミクスの受動性にあるのである。

これを示すため、まず A7) から \$u_d\$ が存在するので、入力偏差を

$$\Delta u_k = u_k - u_d \quad (21)$$

と置こう。そして、A6) の関数 \$F\$ の中で最も簡単と思われる形式

$$u_{k+1}(t) = u_k(t) - \Phi \Delta z_k(t) \quad (22)$$

を考えてみよう。ここに \$\Phi\$ は適当な正定値対称行列とする。式 (22) の両辺から \$u_d\$ を引くと、

$$\Delta u_{k+1} = \Delta u_k - \Phi \Delta z_k \quad (23)$$

になる。そこで、両辺についてそれぞれ \$\Phi^{-1}\$ を介して内積をとると

$$\begin{aligned} \Delta u_{k+1}^T \Phi^{-1} \Delta u_{k+1} &= \Delta u_k^T \Phi^{-1} \Delta u_k \\ &+ \Delta z_k^T \Phi^{-1} \Delta z_k - 2 \Delta z_k^T \Delta u_k \end{aligned} \quad (24)$$

となる。ここで

$$\|\Delta u_k\|^2 = \int_0^T \Delta u_k^T(t) \Phi^{-1} \Delta u_k(t) dt$$

と定義しよう。式 (24) は、

$$\begin{aligned} \|\Delta u_{k+1}\|^2 &= \|\Delta u_k\|^2 + \|\Phi \Delta z_k\|^2 \\ &- 2 \int_0^T \Delta z_k^T(t) \Delta u_k(t) dt \end{aligned} \quad (25)$$

と書き表される。ここに \$\|\cdot\|\$ はヒルベルト空間 \$L^2(0, T)\$ の関数ノルムとみなせる。もし、式 (25) の右辺の積分が正値をとるばかりか、\$\|\Phi \Delta z_k\|^2\$ を相殺して余りある最も望ましい場合を考える。すなわち、ある \$\beta > 0\$ があって、不等式

$$\begin{aligned} &\int_0^t \Delta z_k^T(\tau) \Delta u_k(\tau) d\tau \\ &\geq \frac{1+\beta}{2} \int_0^t \Delta z_k^T(\tau) \Phi \Delta z_k(\tau) d\tau \end{aligned} \quad (26)$$

が成立したとする。これを式 (25) に代入すると

$$\|\Delta u_{k+1}\|^2 \leq \|\Delta u_k\|^2 - \beta \|\Phi \Delta z_k\|^2 \quad (27)$$

を得る。この不等式から次のとこが結論できる、

結論2 (可学習性): もし式 (26) が満足されていれば、偏差 Δz_k が L^2 -ノルムの意味で0にならない限り、入力偏差のノルム $\|\Delta u_k\|^2$ は練習の繰り返しごとに減少する。さらに、

$$\lim_{k \rightarrow \infty} \Delta z_k = 0 \quad (28)$$

である。つまり、練習を積み重ねると、 $\Delta z_k \rightarrow 0$ となる。

もし、 $\Delta z_k \rightarrow 0$ が $\Delta y_k \rightarrow 0$ を意味するなら、これで理想出力 $y_d(t)$ が得られ、制御入力 $u_d(t)$ が練習によって獲得できたことになる。そこで、 Δz をどのように選ぶと不等式 (26) が満足されるかが問題の鍵となる。

いま、 k 回目の練習でロボットのダイナミクスは式 (2) に従うとすれば、理想の軌道を与える式 (3) との偏差は式 (4) で表される、つまり、

$$\{H_0 + H(q_k)\} \Delta \ddot{q}_k + \left\{ B + \frac{1}{2} \dot{H}(q_k) + S(q_k, \dot{q}_k) \right\} \Delta \dot{q}_k + A_1 s(\Delta q_k) + h(\Delta q_k, \Delta \dot{q}_k) = \Delta u_k \quad (29)$$

と表されるとき考えられる。そして、 $\Delta z = \Delta \dot{q} + \alpha s(\Delta q)$ と選んでみよう。このとき、式 (13) と同様に式

$$\int_0^t \Delta z_k^T(\tau) \Delta u_k(\tau) d\tau \geq W(\Delta q_k(t), \Delta \dot{q}_k(t)) - W(\Delta q_k(0), \Delta \dot{q}_k(0)) + \int_0^t V(\Delta q_k(\tau), \Delta \dot{q}_k(\tau)) d\tau \quad (30)$$

が成立する。そこで、

$$\bar{V}(\Delta q, \Delta \dot{q}) = V(\Delta q, \Delta \dot{q}) - \frac{1+\beta}{2} \Delta z^T \Phi \Delta z \quad (31)$$

と定義したとき、 $\bar{V}$ が $\Delta q, \Delta \dot{q}$ について正定になるように B_1, A_1 が取れるかどうか考えてみる。 V が正定になるように A_1, B_1 が選べるなら、 Φ が B_1 くらいでかつ $0 < \beta < 1$ と考えると、 $\bar{V}$ も正定になるように A_1, B_1 を V の正定性が成立するときより少し大きめに取れるはずである。繰り返し学習の場合は、ある定数 $\lambda > 0$ が存在して、形式

$$\lambda W(\Delta q, \Delta \dot{q}) + V(\Delta q, \Delta \dot{q}) \quad (32)$$

の正定性が成立すれば十分なのであるが、ここでは、当分 $\bar{V}$ そのものの正定性を仮定してみよう。このとき、A3) の仮定のもとに、式

$$\begin{aligned} \int_0^T \Delta z_k^T(t) \Delta u_k(t) dt &\geq W(\Delta q_k(T), \Delta \dot{q}_k(T)) \\ &- W(\Delta q_k(0), \Delta \dot{q}_k(0)) + \int_0^T \bar{V}(\Delta q_k(t), \Delta \dot{q}_k(t)) dt \\ &+ \frac{1+\beta}{2} \|\Phi \Delta z_k\|^2 \geq \frac{1+\beta}{2} \|\Phi \Delta z_k\|^2 \end{aligned} \quad (33)$$

が成立する。つまり、式 (33) は式 (26) を意味し、こうして式 (27) と式 (28) が成立することが判る。すなわち、形式

$$\Delta \dot{q}_k + \alpha s(\Delta q_k) = \Delta z_k \quad (34)$$

を微分方程式と見たとき、右辺の強制項が $k \rightarrow \infty$ のとき 0 に L^2 -ノルムの意味で収束するので、解 $\Delta q_k(t)$ は $k \rightarrow \infty$ のとき L^2 -ノルムの意味で 0 に収束することが判る。実際には、 $\Delta \dot{q}_k(t)$ が L^2 -ノルムの意味で $k \rightarrow \infty$ のとき 0 に収束すれば、 $\Delta q_k(t)$ は一様ノルムの意味で 0 に収束する。こうして、次の結論が引き出せた。

結論3 繰り返し学習則 (22) のもとで $\Delta z_k = \Delta \dot{q}_k + \alpha s(\Delta q_k)$ とし (図2参照)、 A_1, B_1 が式 (14), (15), (31) で定義した W と $\bar{V}$ を正定にするように選べたならば、試行回数 k を増加させるとき $q_k(t)$ は理想の軌道 $q_d(t)$ に $t \in [0, T]$ について一様に収束する。

以上はごく最近得られた結果 [1] であるが、これは今までの結果と幾つかの点で異なる。そこで、今までの考察との対比を与えておく。

(1) 学習制御の研究の初期では加速度偏差 ($\Delta \ddot{q}_k$) を使った D 型学習則を展開したが、この場合練習中の位置や速度の軌道の一様有界性は局所的にしか成立しない [2]。

(2) 速度偏差 $\Delta \dot{q}_k$ のみを用いる P 型学習則のとき、軌道の大域的一様有界性が示されたが [3]、その証明では L^2 -ノルムに指数関数の重みが付けられた。その結果、初期化誤差やダイナミクスのゆらぎがあると、学習則に忘却因子を導入しなければならない [4]。

(3) 実験によると [5]、P 型よりも PI 型の学習則の方が良い結果を与えるが、PI 型では大域的な議論が出来ない。

(4) 位置誤差について飽和型のフィードバックを行った PI 型の学習則を使うと、普通には受動性が大域的に成立し、証明のテクニックに指数関数の重み付けを必要としない。したがって、一様有界性や収束性が大域的に成立する。

図2 SP-D方式に基づく繰り返し学習

(5) 理想の速度軌道 $\dot{q}_d$ が高速であって, 式 (15) の V が正定にならなくても, A1) の $T > 0$ が有限なので, 結論 2 はやはり成立する. 何故なら, 式 (30) の代わりに

$$\begin{aligned} \int_0^T e^{-\lambda t} \Delta z_k^T(t) \Delta u_k(t) dt &\geq e^{-\lambda T} W(\Delta q_k(T), \Delta \dot{q}_k(T)) \\ &\quad - W(\Delta q_k(0), \Delta \dot{q}_k(0)) \\ &\quad + \int_0^T [\lambda W(\Delta q_k(t), \Delta \dot{q}_k(t)) + V(\Delta q_k(t), \Delta \dot{q}_k(t))] dt \end{aligned} \quad (35)$$

が成立するからである. このとき, $\lambda > 0$ を適当に大きく取れば, 式 (32) の形式は $\Delta q_k, \Delta \dot{q}_k$ について正定になり, こうして式 (27) がノルム

$$\|\Delta u\|^2 = \int_0^T e^{-\lambda t} \Delta u_k^T(t) \Phi^{-1} u_k(t) dt \quad (36)$$

の意味でやはり成立するからである.

(6) 手先が幾何学的拘束を受ける場合も P 型学習則について考察されているが [5]–[7], 上述のような位置誤差を飽和させた PI 型の学習則も考えられ得ることは容易に推察されよう [1].

(7) ダイナミクスの近似モデルが既知であり, それが利用できれば, 学習の収束性のスピードアップが可能である [8].

4. 初期化を行わない繰り返し制御

理想軌道 $q_d(t)$ が周期 $T > 0$ で周期的な場合, 式 (31) の $V(\Delta q, \Delta \dot{q})$ が $\Delta q, \Delta \dot{q}$ について正定になるように A_1, B_1 が取れば, 初期化なしで位置軌道が $t \rightarrow \infty$ のとき理想の周期軌道に収束することが示せる. このような制御問題を従来, 繰り返し制御法と呼び, 詳しい研究が行われたが [9]–[11], ロボットのような非線形ダイナミクスに基づいて収束性が厳密に論じられたことは一度もなかった. 実際には, ロボットを含む機械系については, 繰り返し制御が成立するための本質は偏差ダイナミクスの受動性にあるのである.

実際, $q_d(t)$ が周期 $T > 0$ で周期的であるとき, $q_k(T)$ は次の周期の初期値 $q_{k+1}(0)$ と見なせるので, 式 (33) の前半の不等式を式 (25) に代入することにより, 式

$$\begin{aligned} \|\Delta u_{k+1}\|^2 + W(\Delta q_{k+1}(0), \Delta \dot{q}_{k+1}(0)) \\ \leq \|\Delta u_k\|^2 + W(\Delta q_k(0), \Delta \dot{q}_k(0)) \\ - \beta \int_0^T \bar{V}(\Delta q_k(t), \Delta \dot{q}_k(t)) dt \end{aligned} \quad (37)$$

を得る. $\bar{V}$ が正定なので, このことは

$$\lim_{k \rightarrow \infty} \int_0^T \bar{V}(\Delta q_k(t), \Delta \dot{q}_k(t)) dt = 0 \quad (38)$$

を意味し, しかも $\bar{V}(\Delta q_k, \Delta \dot{q}_k)$ は $\Delta q_k, \Delta \dot{q}_k$ について 2 次形式的なので, Δq_k と $\Delta \dot{q}_k$ はそれぞれ $L^2[0, T]$ のノルムの意味で $k \rightarrow \infty$ のとき 0 に収束する. $\Delta \dot{q}_k$ の L^2 -ノルムの収束性は Δq_k の一様ノルムの収束性を意味し, 次の結論が示された.

結論 4 初期化をしない式 (22) に基づく繰り返し制御で $\Delta z_k = \Delta \dot{q}_k + \alpha(\Delta q_k)$ とし, A_1, B_1 が結論 2 の仮定を満足するならば, 軌道 $q_k(t) (= q(t + kT))$ は理想の周期軌道 $q_d(t) (= q_d(t + kT))$ に $k \rightarrow \infty$ のとき一様収束する.

なお, 繰り返し制御に関する上記の考え方は, 手先幾何拘束を受けた場合への拡張が可能であるが, ここではこれ以上の深入りはしない. 線形系についてはあったが, 繰り返し制御や学習に受動性が本質的であることを見抜いたのは川村等 [12] が最初である.

5. モデルベース適応制御

偏差ダイナミクスの受動性は適応性の鍵ともなる. よく知られているように [13]–[15], 質量や慣性モーメントの動的パラメータはラグランジュの運動方程式に係数として線形に入る. これらのパラメータ Θ_i の全体をベクトル $\Theta = (\Theta_1, \dots, \Theta_m)^T$ で表すと,

$$\begin{aligned} Y(q, \dot{q}, \ddot{q})\Theta &= \{H_0 + H(q)\}\ddot{q} \\ &\quad + \left\{B_0 + \frac{1}{2}\dot{H}(q) + S(q, \dot{q})\right\}\dot{q} + g(q) \end{aligned} \quad (39)$$

と表される. ここに Y は回帰子 (regressor) と呼ばれるが, Y 中の 2 つの $\dot{q}$ は区別し, 後者の $\dot{q}$ は式 (39) の右辺第 2 項の $\{\}$ の外側の $\dot{q}$ を表す. 実際には Θ は正確には求まらないので, その推定量を $\hat{\Theta}$ で表し, $\hat{\Theta}$ に基づいて計算したモデルを

$$\begin{aligned} Y(q_d, \dot{q}_d, \ddot{q}_d)\hat{\Theta} &= \{\hat{H}_0 + \hat{H}(q_d)\}\ddot{q}_d \\ &\quad + \left\{\hat{B}_0 + \frac{1}{2}\dot{\hat{H}}(q_d) + \hat{S}(q_d, \dot{q}_d)\right\}\dot{q}_d + \hat{g}(q_d) \end{aligned} \quad (40)$$

と表そう. そして適応制御則を

$$u = v - A_1 s(\Delta q) - B_1 \Delta \dot{q} + Y(q_d, \dot{q}_d, \ddot{q}_d)\hat{\Theta} \quad (41)$$

と設定してみる. ここに v は仮りに置いたエキストラ入力で, 実際には $v \equiv 0$ とする. この適応制御則の中の Y は与えられた理想軌道だけに依存するので, あらかじめ計算してメモリーに貯えておくことができる. なお, パラメータ推定量 $\hat{\Theta}(t)$ については後でその更新則を与える. 式 (41) を式 (II-14) に代入すると,

$$\begin{aligned} Y(q, \dot{q}, \ddot{q})\Theta + \{Y(q, \dot{q}, \ddot{q}) - Y(q_d, \dot{q}_d, \ddot{q}_d)\}\Theta \\ - Y(q_d, \dot{q}_d, \ddot{q}_d)\Delta\Theta + B_1 \Delta \dot{q} + A_1 s(\Delta q) = v \end{aligned} \quad (42)$$

となる。ここに $\Delta\theta = \hat{\theta} - \theta$ とおいた。左辺の第2項の $\{ \}$ は式(5)の $h(\Delta q, \Delta \dot{q})$ に一致することに注意されたい。また、左辺第1項と $B_1 \Delta \dot{q}$ と $A_1 s(\Delta q)$ を加えた項は式(4)の左辺の h 以外の総和に一致している。従って、 $\Delta z = \Delta \dot{q} + \alpha s(\Delta q)$ とすると、 Δz と式(42)との内積は

$$\Delta z^T v \geq \frac{d}{dt} \left[W(\Delta q, \Delta \dot{q}) + \frac{1}{2} \Delta \theta^T \Gamma \Delta \theta \right] + V(\Delta q, \Delta \dot{q}) \quad (43)$$

と書ける。ただし、

$$\frac{d}{dt} \Delta \theta = -\Gamma^{-1} Y^T(q_d, \dot{q}_d, \ddot{q}_d) \Delta z \quad (44)$$

と定めるものとする。つまり、パラメータ更新則を次のように定める。

$$\begin{aligned} \hat{\theta}(t) &= \hat{\theta}(0) \\ &\quad - \int_0^t \Gamma^{-1} Y^T(q_d(\tau), \dot{q}_d(\tau), \ddot{q}_d(\tau)) \Delta z(\tau) d\tau \end{aligned} \quad (45)$$

実際、未知パラメータ θ は定数から成るので、式(45)を時刻 t で微分すると式(44)を得る。こうして、次の結論が引き出せた。

結論5 式(41)と(44)に基づく適応制御システムの支配式(42)の入力 v と出力 $\Delta z (= \Delta \dot{q} + \alpha s(\Delta q))$ の対は、 A_1 、 B_1 を適当に取って式(14)と(15)で定義した W と V が正定になれば、受動性を満足する(図3参照)。

この結論から、 $t \rightarrow \infty$ のとき $\Delta q(t) \rightarrow 0$ となることが導ける。実際、式(42)において $v(t) \equiv 0$ のとき、式(43)から

$$\frac{d}{dt} \left[W(\Delta q, \Delta \dot{q}) + \frac{1}{2} \Delta \theta^T \Gamma \Delta \theta \right] \leq -V(\Delta q, \Delta \dot{q}) \quad (46)$$

図3 回帰子をオフライン計算した時のモデルベース適応制御系。このとき、入出力対 $\{v, \Delta z\}$ は受動性を満足する。

となる。 W と V がともに $\Delta q, \Delta \dot{q}$ に正定なので、 Δq と $\Delta \dot{q}$ はともに $L^2(0, \infty)$ に属し、また、ともに有界かつ一様連続になるので、

$$\lim_{t \rightarrow \infty} \Delta q(t) = 0 \quad \text{and} \quad \lim_{t \rightarrow \infty} \Delta \dot{q}(t) = 0 \quad (47)$$

である[2]。こうして、次の結論が示された。

結論6 結論4と同じ前提条件のもとに、もし $v(t) \equiv 0$ ならば、式(41)と(44)に基づく適応制御則について、 $t \rightarrow \infty$ のとき $q(t)$ と $\dot{q}(t)$ はそれぞれ理想の $q_d(t)$ と $\dot{q}_d(t)$ に収束する。

以上の議論は Slotine and Li[16] の方法を改良し、Sadeh and Horowitz[17] の議論を明解にしている。しかも、回帰子はオンライン計算する必要がなく、また、滑り面 (sliding surface) を導入する必要も全くないことに注意されたい。なお、前々節にも述べたように、運動がきわめて高速のとき V を正定にするような B_1 、 A_1 が取れない恐れがある。その時は、式(41)に代わって制御則を

$$u = -A_1 s(\Delta q) - B_1 \Delta \dot{q} + Y(q, \dot{q}, \ddot{q}_r) \hat{\theta}$$

とする必要がある。ここに $\dot{q}_r = \dot{q}_d - \Lambda \Delta q$ とし、 $S = \dot{q} - \dot{q}_r = \Delta \dot{q} + \Lambda \Delta q$ が滑り面を構成することになる。これが、本来、Slotine and Li によって導入された考え方である。この場合、回帰子の実時間オンライン計算は回避できない。なお、モデルベース適応制御に関する詳細な実験は Whitcomb 等[18] によって報告されているので、参照されたい。その他、パラメータ推定量 $\hat{\theta}(t)$ の収束性の吟味も重要な問題であるが、ここでは割愛する。

6. お わ り に

ロボットの腕でも、適当なサーボループを構築すれば、簡単な学習則や適応制御則の下で、習熟や適応が進む。ここでは数学的に複雑な議論を展開したが、それは習熟や適応が確実に進むことを理論的に示すためであったことに注意したい。特に、学習則そのものは非常に簡単である。動物の行う運動の習熟プロセスも複雑ではないのではなかろうか。ただし、ここで述べた考え方以外の別の規範や異なる素過程もあり得よう。例えば、疲労を最も少なくするように習熟する規範があるかもしれない。現実の習熟プロセスは、これら幾つかの規範の重畳の下で進むかもしれない。

次回は、手先拘束の場合の軌道追従の問題を講述した後、全体論に帰り非線形回路理論を通して「技量の力学的理解」が進められ得るかどうか議論してみたい。最後に、いくつかの重要な研究課題を論じて、講義を締めくくりたい。

参 考 文 献

- [1] S. Arimoto et al.: "Learnability and adaptability in constrained dynamics," Proc. of 1993 IEEE/Nagoya University WWW on Learning and Adaptive System, Nagoya, pp.69-76, 1993.

- [2] 有本卓: “システムと制御の数理”, 岩波講座, 応用数学, 岩波書店, 1993.
- [3] S. Arimoto: “Learning control theory for robot motion,” Int. J. Adaptive Control and Signal Processing, vol.4, pp.543–564, 1990.
- [4] S. Arimoto et al.: “Robustness of P-type learning with a forgetting factor for robotic motions,” Proc. of 29th IEEE Conference on Decision and Control, Hawaii, pp.2640–2645, 1990.
- [5] 南條義人, 有本卓: “ロボット軌道追従のための学習制御系の設計指針”, 日本ロボット学会誌, vol.11, no.7, pp.1047–1055, 1993.
- [6] 浪花智英, 有本卓: “幾何学的拘束のあるロボットマニピュレータの学習制”, 計測自動制御学会論文集, vol.29, no.4, pp.411–418, 1993.
- [7] T. Naniwa and S. Arimoto: “Learning control for geometrically constrained robot manipulators,” Proc. of IROS'93, Yokohama, pp.746–753, 1993.
- [8] 浪花智英, 有本卓: “手先が幾何学的な拘束を受けるマニピュレータの学習制御 (手先力の収束性と計算機シミュレーションによる検証)”, 日本ロボット学会誌, vol.11, no.7, pp.1056–1065, 1993.
- [9] 中野道雄, 井上真, 山本裕, 原辰次: “繰り返し制御”, 計測自動制御学会, 1989.
- [10] S. Hara, Y. Yamamoto, T. Omata, and M. Nakano: “Repetitive control system: A new type servo system for periodic exogenous signals,” IEEE Trans. on Automatic Control, vol.33, pp.659–668, 1988.
- [11] R. Horowitz: “Learning control of robot manipulators,” ASME J. of Dynamic Systems, Measurement, and Control, vol.115, pp.402–411, 1993.
- [12] 川村貞夫, 宮崎文夫, 有本卓: “学習制御方式のシステム論的考察”, 計測自動制御学会論文集, vol.21, no.5, pp.445–450, 1985.
- [13] H. Mayeda et al.: “A new identification method for serial manipulator arm,” Proc. 9th IFAC World Congress, Budapest, Hungary, pp.2429–2434, 1984.
- [14] H. Kawasaki and K. Nishimura: “Terminal-link parameter estimation and trajectory control of robot manipulators,” Preprints of Symp. on Robot Control '85, Barcelona, Spain, pp.213–218, 1985.
- [15] J.J. Craig: “Adaptive Control of Mechanical Manipulators,” Addison-Wesley, New York, 1988.
- [16] J.J. Slotine, and W. Li: “On the adaptive control of robot manipulators,” International J. of Robotics Research, vol.6, pp.49–59, 1987.
- [17] N. Sadeh and R. Horowitz: “Stability and robustness of adaptive controllers for robotic manipulators,” International J. of Robotics Research, vol.9, no.3, pp.74–82, 1990.
- [18] L.L. Whitcomb, A.A. Rizzi, and D.E. Koditschek: “Comparative experiments with a new adaptive controller for robot arms,” IEEE Trans. on Robotics and Automation, vol.9, no.1, pp.59–70, 1993.

有本 卓 (Suguru Arimoto)

1936 年 8 月 3 日生。1959 年京都大学理学部卒業。沖電気工業 KK, 東京大学工学部助手, 講師を経て 1968 年大阪大学基礎工学部助教授, 1973 年同教授, 1988 年東京大学工学部教授となり現在に至る。工学博士。ロボティクス, 制御理論, 信号処理, 情報理論の研究に従事。日本機械学会, 情報処理学会, 計測自動制御学会などの会員。IEEE の Fellow 会員。(日本ロボット学会正会員)