数理解析研究所講究録
90巻 1970年 162-170

制御過程と学習過程

九大理情報研　加納省吾

1. 序

学習過程の数学的モデルには線型モデル、非線型モデル、吸収マルコフモデル、エルゴードモデル 等が考えられており、モデル過程の収斂、実験データとの比較なども研究されている。ここでは線型モデルを取り上げ、これを制御過程の応用として表現することを考える。パターン認識と学習過程との間の関係や、確率近似理論と制御過程との間の差異についても、学習過程の構造をはっきりさせることにより明らかになる点が多いと思われる。

2. 学習過程

はじめに学習過程の2つの線型モデルについて説明し、後でこれらを制御過程で表現する。

2.1. 簡単な線型モデル（実験者統制事象の場合）

代表的な実例としては、ねずみの単純T一迷路、訂正法に

よる確率学習や、左右の電球の点滅を左右のキー押しによって学習する確率学習がある。

（記法）

$$X_n = \begin{cases} x_{n1} & \text{，第}n\text{試行の反応の値が右曲り又は右押しのとき（正反応とする）} \\ x_{n2} & \text{，第}n\text{試行の反応の値が左曲り又は左押しのとき（誤反応とする）} \end{cases}$$

$p_n = P\{X_n = x_{n1}\}$ ，第 n 試行で x_{n1} が起る確率（正反応確率という）

$$Y_n = \begin{cases} y_{n1} & , \quad x_{n1} \text{ に対する強化} \\ y_{n2} & , \quad x_{n2} \text{ に対する強化} \end{cases}$$

$\pi_i = P\{Y_n = y_{ni}\}$ ，$i = 1, 2$，強化確率（X_n に独立とする）

$Y^{(n)} = (Y_1, Y_2, \cdots, Y_n)$ ，強化系列（$X_1, \cdots, X_n$ に独立な定常過程とする）

$P\{X_{n+1} = x_{n+1,1} \mid y^{(n)}\}$ ，強化系列が与えられたときの条件つき正反応確率

$y^{(n)}$ ，$Y^{(n)}$ の実現値

第 n 回までの試行と強化（$X_1, Y_1, X_2, Y_2, \cdots, X_n, Y_n$）が行なわれたときの制御は $P\{X_n = x_{n1} \mid y^{(n-1)}\}$ に対してつぎのように行なわれる。

$n = 1$ のとき

$$\begin{cases} P\{X_2 = x_{21} \mid y_{11}\} = \alpha p_1 + (1-\alpha) & , \\ P\{X_2 = x_{21} \mid y_{12}\} = \alpha p_1 & , \quad (0 \leqq \alpha < 1) \end{cases}$$

一般に n のとき

$$\begin{cases} P\{X_{n+1}=x_{n+1,1} \mid y_{n1}, y^{(n-1)}\} = \alpha P\{X_n = x_{n1} \mid y^{(n-1)}\} + (1-\alpha) \\ P\{X_{n+1}=x_{n+1,1} \mid y_{n2}, y^{(n-1)}\} = \alpha P\{X_n = x_{n1} \mid y^{(n-1)}\} \end{cases}$$

いま $y_{n1}=1$, $y_{n2}=0$ とすれば

$$(2,1) \quad P\{X_{n+1}=x_{n+1,1} \mid y_{ni}, y^{(n-1)}\} = \alpha P\{X_n=x_{n1} \mid y^{(n-1)}\} + (1-\alpha) y_{ni}, \quad i=1,2.$$

このようにすると、条件つき正反応確率 $P\{X_n=x_{n1} \mid y^{(n-1)}\}$ の期待値は

$$E\{P(X_n=x_{n1} \mid y^{(n-1)})\} = \sum_{y^{(n-1)}} P\{X_n=x_{n1} \mid y^{(n-1)}\} P\{Y^{(n-1)}=y^{(n-1)}\}$$

$$= \sum_{y^{(n-1)}} P\{X_n=x_{n1}, Y^{(n-1)}=y^{(n-1)}\} = P\{X_n=x_{n1}\} = p_n$$

よって (2,1) から

$$p_{n+1} = \alpha p_n + (1-\alpha)\pi_1 = \alpha^n p_1 + \pi_1(1-\alpha^n) \longrightarrow \pi_1$$

即ち平均値は学習目的である π_1 へ収斂する。

2.2. 距離減少作用素モデル（被験体・実験者統制事象の場合）

2.1の線型モデルでは条件つき正反応確率の制御を決定する事象は実験者の与える強化 Y_n であった。この Y_n が反応 X_n に依存するとモデルは一層複雑になる。NORMAN(1968)はこのような場合を、距離減少モデルの概念を用いて一般的にとり扱っている。その結果を距離減少作用素モデルについて述べる。

$Y_n = y_{n,ij}$, x_{ni} の強化 $i,j=1,2$.

$$\pi_{ij} = P\{Y_n = y_{n,ij} \mid x_{ni}\}, \quad i,j=1,2, \quad \pi_{i1}+\pi_{i2}=1.$$

$$(X_1, Y_1, \cdots, X_n, Y_n) = Z^{(n)}$$

制御は4作用素6パラメーターで

$$P\{X_{n+1}=x_{n+1,1} \mid x_{ni}, y_{n,ij}, z^{(n-1)}\} = (1-\alpha_{ij})P\{X_n = x_{n1} \mid z^{(n-1)}\} + \alpha_{ij}\delta_{j1}$$

$$i,j=1,2.$$

また、反応と強化の確率は

$$P\{X_{n+1}=x_{n+1,i}, Y_{n+1}=y_{n+1,ij} \mid z^{(n)}\} = \{P(X_n = x_{n1} \mid z^{(n-1)})\delta_{i1} +$$

$$+ (1-P(X_n = x_{n1} \mid z^{(n-1)}))\delta_{i2}\}\pi_{ij}, \quad i,j=1,2.$$

以上から パラメーターは α_{ij}, $i,j=1,2$, π_{11}, π_{22} の6コである。制御の形は n に関係しないので、条件つき正反応確率を P_n としてつぎのようにかく。

$$(2.2) \qquad P_{n+1} = f_{ij}(P_n) = (1-\alpha_{ij})P_n + \alpha_{ij}\delta_{j1}, \quad i,j=1,2.$$

また、$X_n = x_{ni}$ と $Y_n = y_{n,ij}$ との同時確率をこの P_n を用いて

$$\varphi_{ij}(P_n) = \{P_n\delta_{i1} + (1-P_n)\delta_{i2}\}\pi_{ij}, \quad i,j=1,2.$$

$0 \leqq P_n \leqq 1$ であるから、状態空間を $S=[0,1]$ とする。

LEMMA 2.1.

$$\forall i \in \{1,2\}, \ \exists j_i \in \{1,2\}, \ \alpha_{ij_i} > 0, \ \pi_{ij_i} > 0$$

ならば

$$1) \qquad \sup_{p \neq p'} \frac{|f_{ij}(p) - f_{ij}(p')|}{|p-p'|} < 1,$$

(2) $\varphi_{ij}(p) > 0$

(3) $$\sup_{p \neq p'} \frac{|\varphi_{ij}(p) - \varphi_{ij}(p')|}{|p - p'|} < \infty$$

このようなモデルを距離減少作用素モデルという。

THEOREM 2.1.

任意の非周期的距離減少作用素モデルに対しては確率1で $P_n \longrightarrow P_\infty$

また、実数 $0<C<\infty, 0<\alpha<1$ が存在して

$$\| E\{P_n^\nu\} - E\{P_\infty^\nu\} \| \leqq C(\nu+1)\alpha^n \quad , \quad \nu \geqq 1$$

n は正整数

こゝに

$$\| f(p) \| = \sup_{p \in S} |f(p)| + \sup_{p \neq p'} \frac{|f(p) - f(p')|}{|p - p'|}$$

3. 制御過程の学習過程への応用

3.1. 線型制御過程

$U = (U_1, U_2, \cdots)$ ，被制御過程

$A = (\xi_1, \xi_2, \cdots)$ ，目標過程

$$C = \begin{pmatrix} C_{11} & 0 & 0 & 0 & \cdots \\ C_{21} & C_{22} & 0 & 0 & \cdots \\ C_{31} & C_{32} & C_{33} & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}$$ ，制御行列（C_{ij}；確率変数）

$V = (V_1, V_2, \cdots)$ ，制御過程

$$U_1 = u_1 \Rightarrow V_{2i} = U_i + c_{11}(\xi_1 - u_1), \quad i \geqq 2$$

$$V_{22} = v_2 \Rightarrow V_{3i} = V_{2i} + c_{21}(\xi_1 - u_1) + c_{22}(\xi_2 - v_2), \quad i \geqq 3$$

$$V_{33} = v_3 \Rightarrow V_{4i} = V_{3i} + c_{31}(\xi_1 - u_1) + c_{32}(\xi_2 - v_2) + c_{33}(\xi_3 - v_3), \quad i \geqq 4$$

$$\cdots \qquad \cdots \qquad \cdots$$

$V = (U_1, V_{22}, V_{33}, \cdots)$ とする。

THEOREM 3.1.

$$V_{m+1} = U_{m+1} + \sum_{k=1}^{m} e_{m,k}(\xi_k - U_k)$$

こゝに e_{mk}, $k = 1, 2, \cdots, m$ はつぎで与えられる。

$$\begin{cases} e_{mm} = c_{mm} \\ e_{m,m-1} = \sum_{k=m-1}^{m} c_{k,m-1} - e_{mm} c_{m-1,m-1} \\ e_{m,m-2} = \sum_{k=m-2}^{m} c_{k,m-2} - e_{mm} \sum_{k=m-2}^{m-1} c_{k,m-2} - e_{m,m-1} c_{m-2,m-2} \\ \cdots \qquad \cdots \qquad \cdots \\ e_{m,1} = \sum_{k=1}^{m} c_{k,1} - e_{mm} \sum_{k=1}^{m-1} c_{k,1} - \cdots - e_{m2} c_{11}. \end{cases}$$

THEOREM 3.2.

(i) $U - A$ ，L_1 有界マーチンゲール

(ii) $E\{c_{mk} \mid c_{ij},\ i = 1, 2, \cdots, m-1,\ j = 1, 2, \cdots, i\} = c_{m-1,k}$, $k = 1, 2, \cdots, m-1$.

(iii) $c_{ii} = 1$, $i = 1, 2, \cdots$

(iv) $E\{(c_{mk} - c_{m-1,k})^2 \mid c_{ij},\ i = 1, \cdots, m-1,\ j = 1, \cdots, i\} = O(m^{-(4+\delta)})$

$\delta > 0$

とすると　$V - A$　は確率1で収斂する。

3．2．学習過程への応用

学習過程の線型モデルに対しては　(2.1) 式を

$$P_{n+1} = \alpha P_n + (1-\alpha) Y_n = P_n + (1-\alpha)(Y_n - P_n)$$

とかくと、これは線型制御でつぎのような場合に当る。

$$U = (p_1, p_1, \cdots)$$

$$A = (Y_1, Y_2, \cdots)$$

$$C_{ij} = \begin{cases} 1-\alpha & i=j=1,2,\cdots \\ 0 & i \neq j \end{cases}$$

$$V = (p_1, P_2, P_3, \cdots)$$

この場合　$1-\alpha$　は定数で　THEOREM 3.1. の特殊の場合となる。THEOREM 3.1. の具体的なものについては文献(5)参照．

つぎに、距離減少4作用素モデルに対しては (2.2) 式を

$$P_{n+1} = P_n + \alpha_{ij} (\delta_{ij} - P_n)$$　とかくとこれは

$$U = (p_1, p_1, \cdots)$$

$$A = (\delta_{ij}, \delta_{ij}, \cdots)$$

$$C_{ij} = \begin{cases} \alpha & i=j=1,2,\cdots \\ 0 & \text{その他のとき} \end{cases}$$

とした場合に当る。ここに　$P\{\alpha = \alpha_{ij} \mid x_{ni}\} = \pi_{ij}$.

4. 学習過程と制御過程との関係

両過程を比較してつぎのような点に気がつく。

(1). 学習過程では一般にAはP_nに依存した確率過程となる。このことは制御過程でもこのような拡張が可能であることを示している。これは多次元の制御過程の特殊の場合と考えられる。多次元制御過程については文献(6)参照。

(2). 学習過程ではAは1次元よりはむしろ多次元として考えた方がよい。これは統制事象としていろいろなものを考えた方が学習が効果的であることによる。このような場合でもAの確率過程の性質としては多次元マーケンゲールで十分と思われる。

(3). 学習の目標は一応Aによつて与えられるが、この中には学習を効果的にするための変数も含まれており、これらを総合した学習目標が考えられなければならない。制御過程の制御目標としても、このような目的達成を効果的にする変数が考えられてもよい。

(4). 今までの学習過程は記憶のない場合(毎回、過去のデータを用いない)のみを扱っているが、制御過程は記憶のある場合の学習過程とみなすことができる。

参考文献

(1) ATKINSON, R. C., BOWER, G. H. & CROTHERS, E. J. (1965): An Introduction to Mathematical Learning Theory. Wiley.

(2) NORMAN, M. F. : Some convergence theorem for stochastic learning Models with distance diminishing operators, J.M.P. Vol. 5, No. 1. (1968)

(3) 印東太郎 編 数理心理学 東大出版会 (1969)

(4) KITAGAWA, T. : Successive processes of statistical controls (3), Mem. Fac. Sci., Kyusyu Univ. Ser. A, Vol. 14, No. 1 (1960)

(5) KANÔ, S. : Linear controlled stochastic processes, Sci. Rep. Kagoshima Univ., No. 8 (1959)

(6) KANÔ, S. : Multi-dimensional linear controlled stochastic processes with discrete parameter, Sci. Rep. Kagoshima Univ., No. 10, (1961)

(7) KANÔ, S. : Martingale transforms and linear controlled stochastic processes, Res. Rep. of Res. Inst. of Fund. Inf. Sci., No. 5 (1969)