構造未知のシステムに対する シミュレーション・アプローチ

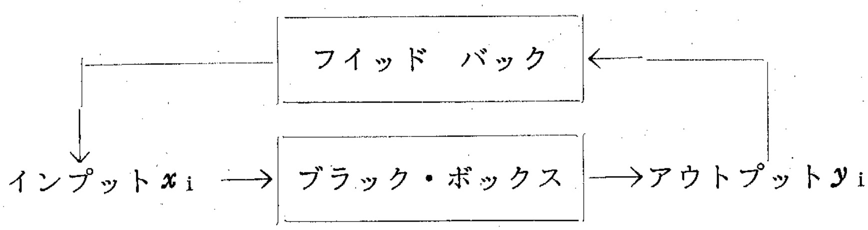
伊　藤　駒　之

1　序

ここで言及するシステムの構造未知部分はブラック・ボックス（Black. Box）とする。ゆえに，システムは，インプット（Input），ブラック・ボックス，アウトプット（Output），フィードバック（Feedback）の各要素から構成される。つぎのような仮定をおく。

アウトプット y_i（$i=1$，…，N）はインプット x_i をブラック・ボックスが変換した結果であり，一つのインプット x_i に対してアウトプット y_i が一意的に対応し，アウトプット y_i は観測可能であり，このアウトプットを評価するスカラー関数 $U(y_i)$ は与えられている。ブラック・ボックスは未知で，静態的すなわち時間の経過と共に変遷することはない。また確率的要因も含まない。すなわち決定論的（deterministic）であるとする。

上記の各システム要素を図式に表現すると(1)図のようになる。



（1）図

我々が問題とするのは，(1)図に示すフイードバックに関連するコントロールである。フイードバックの目的であるコントロールを行なうプロセスは試行錯誤的方法（すなわち heuristic approach）である。ゆえに，このフイードバ

ック・コントロールのプロセスは学習過程でもある。我々は，このような観点から，一つの学習モデルを構成することになる。

我々のモデルは，基本的には，山登り（hill-climbing）である。すなわち，大きい効用があると思われる領域を探索することに比重をかけ，小さい効用があると思われる領域に小さい比重をかける。インプット x_i に対応している加重 u_i は確率とみなすことが出来るように規準化（$\sum u_i = 1$）している。

つぎのような目的を設定する。

「 $\underset{1 \leq i \leq N}{\mathrm{Max}}\, U(y_i)$ 」

この目的は仮に設定しただけである。ゆえに，我々のモデルの利用しうる領域はこの目的に限定されない。

2 モデルの基本的構造

未知のブラック・ボックスに対して，フイードバック・コントロールがいかに適応していくかを示すために，モデルの基本的構造を叙述する。Bellman の確率的学習モデルの表現を借りることにする[①]。

各段階において，N 個の選択対象 $x_1, \cdots, x_N$が選択可能である多段階過程（multistage process）を考え，選択対象 x_i を選択する確率を P_i とする，初期確率分布 $F_1 = (P_1, \cdots, P_N)$ は任意のものとする。

そして，ある選択対象 x_i が，初期確率分布 F_1 に従って，選択されると，この x_i に依存して，ある結果が生じる。この結果についての情報がつぎの対象選択に変化をもたらす。この対象選択の変化を示すために，初期確率分布 F_1 を新しい確率分布 F_2 に変換する。ここで

$$F_2 = T_2(F_1) = [T_{21}(F_1), \cdots, T_{2N}(F_1)] \quad -2.1$$

である。F_2 は確率分布を表現しているものであるから，$T_2(F_1)$ の各要素はつぎのような条件を満たさなければならない。

(a) $T_{2j}(F_1) \geqq 0$

(b) $\sum_{j=1}^{N} T_{2j}(F_1) = 1 \qquad -2.2$

このような手続で，プロセスは進行することになる。

以下，このプロセスの詳細な手続を記述する。

① R.Bellman, "*Adaptive Control Processes*," Princeton University Press, 1961, Chapter XIII, 13.2 A stochastic learning model, 参照.

3 適 応 過 程

インプット x_1, ……, x_N はスカラーで，離散的とする。のちに，インプット領域の多元化，連続化への拡張を行なう。

各段階において，N個のインプット x_1, ……, x_Nが選択対象である。

第一段階では，我々は，ある確率分布 $F_1=(P_{11}, \cdots, P_{1N})$ でもって，インプットを選択する。P_{1i} は x_i を選択する確率である。さて，2つのインプット $x_{i(1)}$, $x_{i(2)}$ が選択されたとする。$x_{i(1)} \rightarrow y_{i(1)} \rightarrow U(y_{i(1)})$, $x_{i(2)} \rightarrow y_{i(2)} \rightarrow U(y_{i(2)})$ となり，$U(y_{i(1)})$, $U(y_{i(2)})$ が決定される。矢印→は因果関係を示す。

第二段階。　$i_{(1)} < i_{(2)}$とする。

$\bar{i}_{(2)} = \dfrac{i_{(1)}+i_{(2)}}{2}$，$\bar{i}_{(2)}$ は $\dfrac{i_{(1)}+i_{(2)}}{2}$ より大きくない最大の整数を示す。

このような前提で，インプット領域を $\bar{i}_{(2)}$ なる境界にて分割する。すなわち

$$(x_1, \cdots, x_{\bar{i}(2)-1}), \quad (x_{\bar{i}(2)}, \cdots, x_N)$$

いま，

$$A_1=\max\{|U(y_{i(1)})|, |U(y_{i(2)})|\}$$

$$B_1=\min\{|U(y_{i(1)})|, |U(y_{i(2)})|\}$$

$$W_{i(1),1}=(U(y_{i(1)})+2A_1)(A_1-B_1)^{-1}$$

ただし　$A_1-B_1 \neq 0$

$$W_{i(2),1}=(U(y_{i(2)})+2A_1)(A_1-B_1)^{-1}$$

$$\alpha_1=\sum_{k=1}^{\bar{i}(2)-1} W_{i(1),1}+\sum_{k=\bar{i}(2)}^{N} W_{i(2),1}$$

とおく。

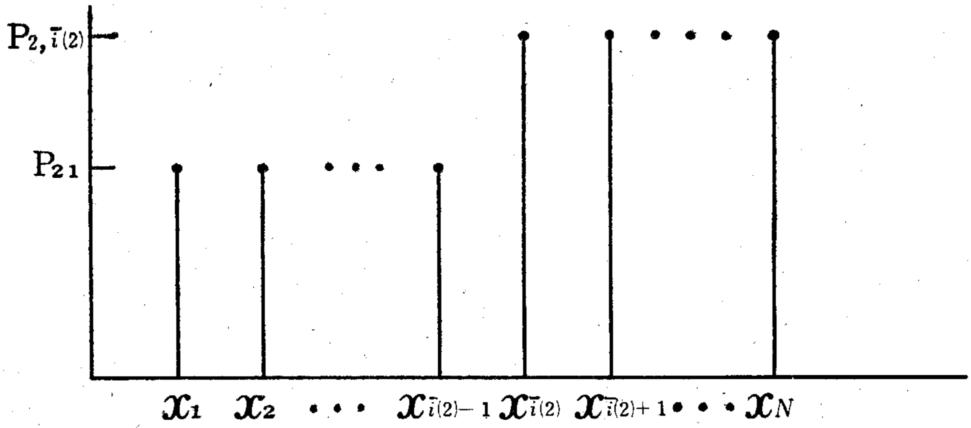
$$P_{2\ell}=\begin{cases}\dfrac{W_{i(1),1}}{\alpha_1} & \ell=1,\cdots,(\bar{i}_{(2)}-1)\\[2ex] \dfrac{W_{i(2),1}}{\alpha_1} & \ell=\bar{i}_{(2)},\cdots,N\end{cases}$$

から，インプットを選択する確率分布 F_2 は

$$F_2=(P_{21},\cdots,P_{2,(\bar{i}(2)-1)},P_{2\bar{i}(2)},\cdots,P_{2N})$$

と定義する。$P_{2\ell}$ は第2段階で x_ℓ を選択する確率である。ただし $A_1-B_1=0$ なら F_1 を使用する。

いま，得られたインプット選択の確率分布 F_2 を図で表現すると，(2)図のようになる。



(2) 図

つぎに，$x_{i(3)}$ が選択されたとする。そうすると，$x_{i(3)}\to y_{i(3)}\to U(y_{i(3)})$ となり，$U(y_{i(3)})$ が定まる。

第3段階。$i_{(1)}<i_{(3)}<i_{(2)}$ とする。

$$\bar{i}_{(3)}=\left[\frac{i_{(1)}+i_{(3)}}{2}\right]$$

$$\bar{i}_{(2)}=\left[\frac{i_{(3)}+i_{(2)}}{2}\right]$$

このような前提で，インプット領域を $\bar{i}_{(2)}$，$\bar{i}_{(3)}$ なる2つの境界にて3分割する。すなわち，

$$(x_1,\cdots,x_{\bar{i}(3)-1}),\ (x_{\bar{i}(3)},\cdots,x_{\bar{i}(2)-1}),\ (x_{\bar{i}(2)},\cdots,x_N)。$$

いま，

$$A_2=\max\{|U(y_{i(1)})|,|U(y_{i(2)})|,|U(y_{i(3)})|\}$$

$$B_2=\min\{|U(y_{i(1)})|,|U(y_{i(2)})|,|U(y_{i(3)})|\}$$

$$w_{i(1),2}=(U(y_{i(1)})+2A_2)(A_2-B_2)^{-1}$$

$$w_{i(3),2}=(U(y_{i(3)})+2A_2)(A_2-B_2)^{-1}$$

$$w_{i(2),2}=(U(y_{i(2)})+2A_2)(A_2-B_2)^{-1}$$

$$\alpha_2=\sum_{k=1}^{\bar{i}(3)-1}w_{i(1),2}+\sum_{k=\bar{i}(3)}^{\bar{i}(2)-1}w_{i(3),2}+\sum_{k=\bar{i}(2)}^{N}w_{i(2),2}$$

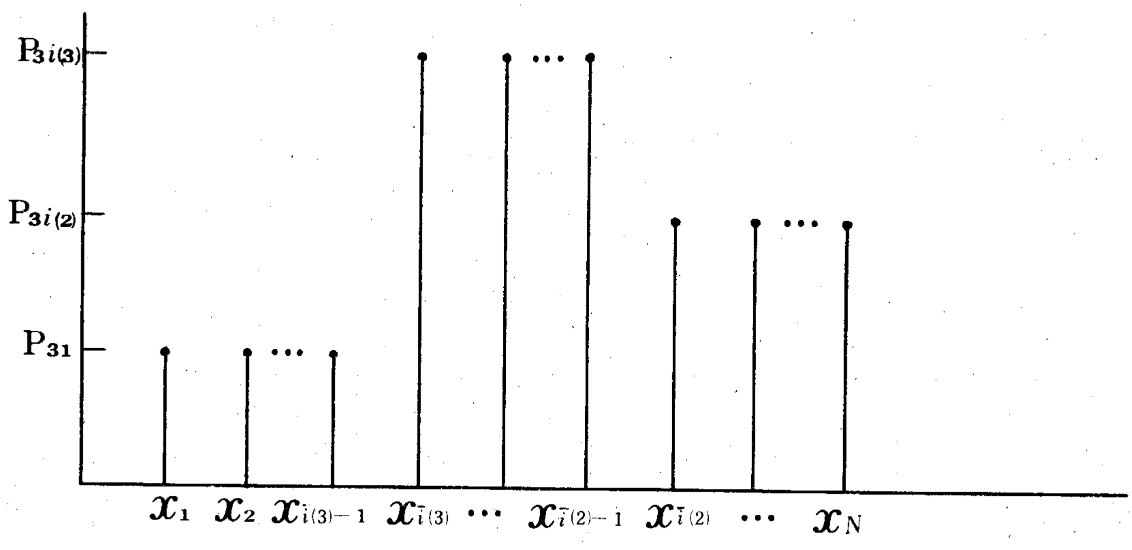
とおく。

$$P_{3\ell}=\begin{cases}\dfrac{w_{i(1),2}}{\alpha_2} & \ell=1,\cdots,(\bar{i}_{(3)}-1)\\[2ex] \dfrac{w_{i(3),2}}{\alpha_2} & \ell=\bar{i}_{(3)},\cdots,(\bar{i}_{(2)}-1)\\[2ex] \dfrac{w_{i(2),2}}{\alpha_2} & \ell=\bar{i}_{(2)},\cdots,N\end{cases}$$

から，インプットを選択する確定分布 F_3 は

$$F_3=(P_{31},\cdots,P_{3,(\bar{i}(3)-1)},P_{3\bar{i}(3)},\cdots,P_{3,(\bar{i}(2)-1)},P_{3\bar{i}(2)},\cdots,P_{3N})$$

と定義する。ここで，$P_{3\ell}$ は第３段階で x_ℓ を選択する確率である。確率分布 F_3 を図で表現すると，(3)図のようになる。



(3)　図

この手続を，各段階において，くり返すことになる。

4　インプット領域の多元化

インプット領域（または定義域）がk次元であるケースにおけるインプット選択の確率分布を構成することにする。

一般に，インプット領域がk次元化した状態においては，k次元ユークリッド空間の一点が一つのインプットとなる。ここでは，離散的な点だけから構成されているケースに限定し，インプットの個数，すなわち，選択対象はN個とする。

各段階における対象選択分布の構成はインプット領域の分割と分割された領域に対するウェイトずけにより解決される。

さて，インプット領域の分割手続を示すことにする。インプット領域の分割手続としては種々考えられるが，連続量の多元化との一貫性を維持するためと，単純性のために，つぎのような手続をとることにする。

初期分布としてある対象選択分布をもつ　いま，インプット$\boldsymbol{x}_i$，$\boldsymbol{y}_i$が選択されたとする。ここで，$\boldsymbol{x}_i$，$\boldsymbol{y}_i$は

$$\boldsymbol{x}_i=\begin{pmatrix} x_{1i} \\ \vdots \\ x_{ki} \end{pmatrix}, \qquad \boldsymbol{x}_j=\begin{pmatrix} x_{1j} \\ \vdots \\ x_{kj} \end{pmatrix}$$

なるk次元ベクトルとする。そうすると，

$$\boldsymbol{x}_i \to \boldsymbol{y}_i \to U(\boldsymbol{y}_i), \quad \boldsymbol{x}_j \to \boldsymbol{y}_j \to U(\boldsymbol{y}_j)$$

となり，$U(\boldsymbol{y}_i)$，$U(\boldsymbol{y}_j)$が決定される。ここで，$\boldsymbol{y}_i$，$\boldsymbol{y}_j$はベクトルまたはスカラーであり，$U(\boldsymbol{y}_i)$，$U(\boldsymbol{y}_j)$は一意的に定まるスカラー関数である。

$\boldsymbol{x}_i$，$\boldsymbol{x}_j$の第1要素x_{1i}，x_{1j}をとりあげて考察する。第1番目の座標軸において，x_{1i}，x_{1j}はN個の選択対象中$i_{(1)}$番，$j_{(j)}$番目に存在し，

$$\bar{j}_{(1)}=\left[\frac{i_{(1)}+j_{(1)}}{2}\right], \quad \bar{j}_{(1)}\text{は}\frac{i_{(1)}+j_{(1)}}{2}\ \text{より大きくない最大の整数。}$$

とする。このような前提で第1番目の座標軸を$\bar{j}_{(1)}$にて分割する。このような分割法を1個の座標軸各々に適用するとき，2^{l}個の互に素なる領域に分割される。

つぎに対象選択分布の構成を考察する。各座標軸において，インプットがスカラーのケースと全く同じ手続でもって，ベクトルの各要素選択の確率を仮に定める。この確率を変換して対象選択の確率を決める。

いま，インプット x_i についてはつぎのように定まったとする。

$$x_i = \begin{pmatrix} x_{1i} \\ \vdots \\ x_{ki} \end{pmatrix} \begin{matrix} \rightarrow \\ \\ \rightarrow \end{matrix} \begin{pmatrix} P_{1i} \\ \vdots \\ P_{ki} \end{pmatrix}$$

x_i を選択する確率 P_i は

$$P_i = \frac{\sum_{d=1}^{k} P_{ji}}{\sum_{i=1}^{N} \sum_{j=1}^{k} P_{ji}}$$

$$= \frac{\sum_{j=1}^{k} P_{ji}}{k}$$

として定める。

このような手続は，情報が x_i，y_i 以外に増えたときも同様に，対象選択の確率分布を構成することを可能ならしめる。

5　インプット領域の連続化

インプット x はスカラーで，$a < x < b$ なる範囲の一点として定義する。

各段階においては，領域〔a，b〕内の各点が選択対象である。

第一段階では，ある確率分布

$$F_1(x) = \int_a^x f(x)\,dx$$

でもってインプットを選択する。そこで，2つのインプット $x_{i(1)}$，$x_{i(2)}$ が選択されたとする。$x_{i(1)} \rightarrow y_{i(1)} \rightarrow U(y_{i(1)})$，$x_{i(2)} \rightarrow y_{i(2)} \rightarrow U(y_{i(2)})$ となり，$U(y_{i(1)})$，$U(y_{i(2)})$ が定まる。

第二段階。離散的ケースの拡張として類似な手続を構成する。

$$\bar{x}_{i(2)}=\frac{x_{i(1)}+x_{i(2)}}{2}, \qquad x_{i(1)}<x_{i(2)}$$

とする。そうすると，インプット領域は$\bar{x}_{i(2)}$なる境界にて分割可能となる。ゆえに

$$[a,\ \bar{x}_{i(2)}],\ [\bar{x}_{i(2)},\ b]$$

とすることができる。

さて，

$$A_1=\max\{|U(y_{i(1)})|,\ |U(y_{i(2)})|\}$$

$$B_1=\min\{|U(y_{i(1)})|,\ |U(y_{i(2)})|\}$$

$$w_{i(1),1}=(U(y_{i(1)})+2A_1)(A_1-B_1)^{-1}$$

ただし $A_1-B_1 \neq 0$

$$w_{i(2),1}=(U(y_{i(2)})+2A_1)(A_1-B_1)^{-1}$$

$$\alpha_1=\int_a^{\bar{x}i(2)} w_{i(1),1}\, dx+\int_{\bar{x}i(2)}^{b} w_{i(2),1}\, dx$$

とおくと，インプットを選択する確率分布 $F_2(x)$ は

$$F_2(x)=\begin{cases}\displaystyle\int_a^x \frac{w_{i(1),1}}{\alpha_1}dx, & a\leqq x\leqq \bar{x}_{i(2)}\\ \displaystyle\int_a^{\bar{x}i(2)} \frac{w_{i(1),1}}{\alpha_1}dx+\int_{\bar{x}i(2)}^{x}\frac{w_{i(2),1}}{\alpha_1}dx, & \bar{x}_{i(2)}\leqq x\leqq b\end{cases}$$

となる。ただし，$A_1-B_1=0$ なら，初期分布 $F_1(x)$ を第２段階でも使用する。

第３段階も，離散的ケースと類似な手続になる。

連続でかつ多元化したケースについて簡単に記述しておく。

インプット領域の分割法は離散的ケースの多元化からあきらかである。そこで分布の構成についてのべる。

離散的ケースにおけると同様に各座標軸においては，ベクトルの各要素選択確率が定まる。

いま，インプットの各要素について，つぎのように定まったとしよう。

$$x=\begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix} \begin{matrix} \rightarrow \\ \\ \rightarrow \end{matrix} \begin{pmatrix} F_{1i}(x_1) \\ \vdots \\ F_{ki}(x_k) \end{pmatrix}$$

第 i 段階での x を選択する確率分布 $F_i(x)$ は

$$F_i(x)=\frac{\sum_{j=1}^{k} F_{ji}(x_j)}{k}$$

と定義する。

6　停　止　法　則

ここでは，つぎの２つの意志決定のうちどちらを選ぶべきかを定める法則を構成することにする。(1)適応過程を停止する，(2)適応過程を継続する。

ランダム・サンプル（x_1, ……, x_n）の順序統計量を $Z_1 \leq Z_2 \leq \cdots\cdots \leq Z_n$ とすれば，Z_n の確率密度 $g(Z_n)$ は

$$g(Z_n)=n[F(Z_n)]^{n-1}f(Z_n)$$

ただし　$F(x)$：母集団分布

$f(x)$：母集団確率密度

となる。

逐次標本抽出法（Sequential Methods）を適用すると，（n＋1）回目に x_{n+1} をサンプリングした後，（$x_1, \cdots, x_{n+1}$）の順序統計量において，x_{n+1} が

$$x_{n+1} \geq Z_n$$

となる確率 P_{n+1} は

$$P_{n+1}=\int_{Z(n)}^{\infty}(n+1)[F(x)]^n f(x)\,dx$$

となる。

いま，第 n 段階において $x_n \rightarrow y_n \rightarrow U(y_n)$ が決定したとする。そこで，

$$U_i=U(y_i),$$

$$U_k = \max_{1 \le i \le n}\{U_i\},$$

$$\hat{U}_n = \alpha U_{n-1} + (1-\alpha)\hat{U}_{n-1}$$

$$0 < \alpha < 1$$

$$\hat{U}_1 = U_1$$

$$S_n^2 = \frac{\sum_{i=1}^{n}(U_i - \hat{U}_i)^2}{n}$$

とおく。ここで，U_n はランダムでかつ

$$f_n(U_n) = \frac{1}{\sqrt{2\pi}S_n} e^{-\frac{(U_n - \hat{U}_n)^2}{2S_n^2}}$$

なる確率密度をもち，U_k はこの $f_n(U_n)$ における n 個のランダム・サンプル中の最大値と仮定する。

そうすると，（n＋1）回目のサンプリングにおいて

$$U_{n+1} \geqq U_k$$

となる確率は，近似的に，

$$Q_{n+1} = \int_{U(k)}^{\infty} (n+1)\Big[F_n(x)\Big]^n f_n(x)\,dx$$

$$F_n(x) = \int_{-\infty}^{x} f_n(x)\,dx$$

となる。ゆえに，第（n＋1）段階において得られる期待値 E_{n+1} は

$$E_{n+1} = \int_{U(k)}^{\infty} x(n+1)\Big[F_n(x)\Big]^n f_n(x)\,dx$$

となる。

（n＋1）段階を実施するに要する費用 L_{n+1} とすれば，

$$L_{n+1} < E_{n+1}$$

である限り適応過程を継続し，

$$L_{n+1} > E_{n+1}$$

になって適応過程を停止することになる。

7　適応過程の収束性

もし値域に上限が存在すれば，この適応過程が上限に収束することは直感的に明きらかだが，論理性のためその証明を示す。

いま

$$\varphi_n = \max\{U(y_1), \cdots, U(y_n)\}$$

$$\beta = \sup_{x_i \varepsilon X} \{U(y_i(x_i))\}$$

X：インプット領域

とおくと

$$\varphi_1 \leq \varphi_2 \leq \cdots \leq \varphi_n \leq \cdots \leq \beta < \infty$$

となり，数列 $\{\varphi_n\}$ は単調増加でかつ上に有界である。ゆえに，

$$\lim_{n\to\infty} \varphi_n = \sup_n \varphi_n = \beta$$

となり，数列 $\{\varphi_n\}$ は収束する。

8　結

我々の手続，すなわち，効用を規準化した加重（$\Sigma U_i = 1$）に従って，つぎにとる戦略を選択する手続は Von Neumann のゲームの解法①と類似している。

Von Neumann の解法では，加重の大きさはある期待水準と比較してどれほど有利かということに依存している。そして，その期待水準以下の戦略は排除される (delete)，すなわち $u_i = 0$ なる加重をおいている。two person zero Sum Game のケースでは，論理的にこれは可能である。ゆえに，我々もブラック・ボックスの構造からサーチすることの不必要な領域を決定することができるなら，そのような領域を無視するような手続をモデルに組み入れることは可能である。例えばユニモーダルなケースがよい例である。

しかし一般的ケースでは，ブラック・ボックスの構造は単調な山が存在することを保証していない。それで，我々は記録 (history) からみて小さな効用

しか存在しないと考えられる領域にも探索の機会を残すことにしている。このことはモデルの効率（例えば収束速度）を落すかもしれないが，〝落し穴〟に落ちる危険を少なくしている。

「虎穴に入らずんば虎児を得ず」という信念からは

$$\overline{u}_i = \frac{u_i^m}{\Sigma u_i^m} \qquad (m \geq 1)$$

なる戦略をとることも可能である。もちろんmは主観的判断に基ずくものになるだろう。このさい，$m=1$とすれば特殊ケースとしての我々の手続が成立する。

確率分布によるインプットの選択は，真の状態を投影する情報をえるためのルールを構成する一手段として適用された。

もし真の状態を投影する情報を得ることが可能なら，その情報を根拠にしてつぎの探索活動を行なうことは合理的な行動と規定することができる。

しかしながら，我々のモデルでは，各段階においては，真の状態を投影する情報を獲得している保証はない。それで，つぎのような仮定におくことにする。

仮定：我々の手続は，各段階の系列において真の状態を投影する情報を獲得している。

我々のモデルは非常に単純な型を示している。ゆえに，モデルを利用することの可能な分野は大層広大なものと信じる。また，各システム要素は工夫をこらすことにより②，モデルはより有効に働くかもしれない。このような方向で，すなわち新しい方法を例示しかつ改善のための基盤を提供するための第一歩として，我々のモデルの詳細を記述した。

① John Von Neumann, "A Numerical Method to Determine Optimum Strategy," *Naval Research Logistics Quarterly*, 1954.

② 我々のモデルの応用例としてつぎのモデルを提供する。

あるインプット選択分布 F_1 と他のインプット選択 F_2 の比較尺度として，それぞれの効用の期待値を使用する。

分布 F_1 でn個のインプットをサンプリングする。そのサンプルに対応するアウト

プットの効用の標本平均を U_1 とする。

分布 F_2 を F_1 から生成された情報により構成する。F_2 に従って，n個のインプットをサンプリングする。そうすると同様に U_2 が定る。

ここで，F_1 による期待値を U_1，F_2による期待値を U_2 とする。

そして

$$U_1 = U_2$$

なる仮説を検定する。ただし対立仮説は $U_1 < U_2$ とする。

そして帰無仮説が有意なら，分布 F_2 が残る（survive）。

さもなければ，F_1 によりもうn個インプットをサンプリングし，このサンプルを前回のn個のサンプルと統合して，2n個のサンプルを新に F_1 による効用の標本平均を計算する。

そして分布 F_2 はそれまでに獲得された全ての情報から構成される。F_2 に従ってn個のインプットをサンプリングし，このn個に対応する効用から，この新しい分布 F_2 による効用の標本平均を計算する。

このようにして，新たに計算された標本平均について，前回と同様の検定を行なう。

このような手続が継続していくことになる。