綜 合 報 告

多段ゲームとダイナミック・プログラミング

坂 口 実*

1. はしがき

ふつうのいわゆる決定過程(decision process)は，その発生の由来や解析的構造はいろいろであるが，1つの重要な特性を共通にもっている．それは，すべての決定(decision)が，目的函数の値を最大にするという目標に向ってなされることである．この章で考える多段(multi-stage)決定過程にはこの目標の単一性はない．そこではある決定は目的函数を最大にしようとし，他の決定は最小にしようとする．

こういう事態は，対抗する2つの主体が，各自の最適の手 (strategy) を求めようとするときに起る．経済的世界におけるいろいろの事例や，偶然と技術による各種のゲームの例は，この章の理論の面白い適用例になる．また，物理的世界でテストや実験に関連して，自然(nature)が，われわれから真実 (truth) をかくそうとする対抗者であるように考えることが都合がよいのである．実験計画は，この自然からわれわれが情報をひき出そうとする game であると考えられる．

近年発展したgameの理論は，この方面の問題を扱う数学理論である．それは，non Neumann の発見した基本的事実——有名な min-max 定理に基礎を置いている．そこでわれわれは多段 game を論ずる前に，先ずこの定理を簡単に説明しておくことにする．

2. O-和 2-人有限 game

2人の player 甲，乙（以後Ⅰ，Ⅱということにする）がいて，Ⅰは $\alpha_1, \cdots\cdots, \alpha_m$ という m 種類の手 (strategy) の中からどれかを選んで出す．同時にⅡは，甲の出した手は全く知らずに $\beta_1, \cdots\cdots, \beta_n$ という n 種類の手の中からどれかを選んで出す．両者の出した手 α_i, β_j を見ている第三者，あるいは審判 (umpire) がいて，甲には，a_{ij}, 乙には b_{ij} を与えるものとする．これらの数はもちろん符号をもっていて，＋ならば入，－ならば出と考えられる．player が（自分を守るために）固定した手を出したくないときには，混合した手 (mixed strategy)

$$p=(p_1, \cdots\cdots, p_m), \qquad p_i \geqq 0, \quad \sum_{i=1}^{m} p_i = 1$$

$$q=(q_1, \cdots\cdots, q_n), \qquad q_j \geqq 0, \quad \sum_{j=1}^{n} q_j = 1$$

* 電気通信大学短大部数学研究室通信科学研究所

を使ってもよいとする．すなわち，たとえば p の意味は手 α_i を確率 p_i で出すのである．このときには，各人の所得の期待値はそれぞれ

$$\sum_{j=1}^{n}\sum_{i=1}^{m}\alpha_{ij}\,p_i\,q_j, \qquad \sum_{j=1}^{n}\sum_{i=1}^{m}b_{ij}\,p_i\,q_j$$

となる．

今，両者が“所得行例”

$$A=(a_{ij}\mid i=1,\cdots\cdots,\ m\ ;\ j=1,\cdots\cdots,\ n)$$
$$B=(b_{ij}\mid i=1,\cdots\cdots,\ m\ ;\ j=1,\cdots\cdots,\ n)$$

の自分のものはもちろん，相手のものをも知っているものとして，自分の所得の期待値を最大にするためには，どんな手を出せばよいか．

最も簡単なのは O-和，game，すなわち

$$a_{ij}+b_{ij}\equiv 0 \tag{1}$$

の場合である．このとき $b_{ij}=-a_{ij}$ であるから，IIから a_{ij} が審判を素通りしてIに渡されることになる．するとIIの出はそのままIの入であり利害は正反対である．

こういう事態のもとでの方針として von Neumann は min-max 原理を考えた．それは自分の最悪の場合を予想して，そのときを最も有利に行動しようというのである．Iの出す手 p がIIに知れてしまう場合（Iにとって最悪）に，II が $\sum_{i,j}a_{ij}\,p_i\,q_j$ を最小にする手 q を出してくることを予想して，Iはこの最小値を最大にする手 $p=p^0$ を出しておく．IIは自分の手 q がIに知れてしまう場合（IIにとって最悪）にIが $\sum_{i,j}a_{ij}\,p_i\,q_j$ を最大にする手 p を出してくることを予想して，この最大値を最小にする手 $q=q^0$ を出しておく．すると任意の手 p, q に対して

$$\left.\begin{aligned} p^0Aq &\geqq \min_q p^0Aq=\max_p\ \min_q\ pAq\equiv v_1\ ^{(*)}\\ pAq^0 &\leqq \max_p pAq^0=\min_q\ \max_p pAq\equiv v_2 \end{aligned}\right\} \tag{2}$$

であるから，相手の手の如何に拘わらずIは v_1 以上の入，IIは v_2 以下の出を期待できるわけである．

von Neumann の min-max 定理というのは

定理 A が有限行列であれば常に，$v_1=v_2$.

したがって (2) より

$$v_1=\min_q\ p^0Aq=p^0Aq^0=\max_p pAq^0=v^2 \tag{3}$$

が成立する．したがって，I, II 両人が共に min-max 原理に依った手 p^0, q^0 をとるときは，相手の手を知らずに自分の手を出すための損（すなわち，それぞれ

$$\max_p\ pAq^0-p^0Aq^0, \qquad p^0Aq^0-\min_q p^9Aq$$

[註] (*) p を行 vector, q を列 vector とし行列記法で表わしてある．以後この記法を用いる．

である）が何れも 0, という意味で，おのずから，合理的に振るまったことになる．最悪の場合に備えるという悲観的 (pessimistic) 立場に立った手であるから，相手の手がもしわかればより有利に振るまえそうであるが，そうでないのである．(3) の p^0, q^0 を各人の最適の手 (optimal strategy)，更にくわしくは，p^0 を I の max-min な手，q^0 を II の min-max な手という．p^0Aq^0 を game の値 (value) という．

(3) は (p^0, q^0) が函数 pAq の鞍部点 (saddle point) になっていることを示す．min-max 定理の $\max_p \min_q = \min_q \max_p$ は鞍部点が存在するための必要充分条件であるので，O-和 2-人有限 game（これを行例 game ということがある）には必らず最適の手と値とが存在することがわかった．前者は一般に無数，後者は常に唯一に存在する．これらを実際に求めることが応用上重要で，種々の方法が考案されている．

3. O-和 2-人 無限 game

行列 game では両 players のとり得る手が何れも有限個であったが，今度は無限に多数ある場合を考える．たとえば，単位正方形の上の連続 game (continuous game on the unit square) は行列 game の直接の拡張である．I，II が区間 [0, 1] から互に独立に，それぞれ，x, y を選ぶ．payoff は II から I へ $K(x, y)$ とする．mixed strategy は [0, 1] の上の確率分布になる．いま，分布函数：

(i) $F(0) = 0$, $F(1) = 1$, [0, 1] で単調非減少

(ii) (0, 1) で右側連続

なる函数 $F(x)$ の全体を D とかくと，やはり

基本定理 $K(x, y)$ が閉正方形 $0 \leqq x, y \leqq 1$ 上で連続であるならば，

$$\max_{F \in D} \min_{G \in D} \int_0^1 \int_0^1 K(x, y)\, dF(x)\ dG(y)$$

および

$$\min_{G \in D} \max_{F \in D} \int_0^1 \int_0^1 K(x, y)\, dF(x)\ dG(y)$$

が存在して，両者は相等しい．

が成立する．すなわち，この連続 game においても，値および最適の手の存在が保証されているわけである．

基本定理に関連して注意を2つしておく．先ず，$dF(x)$, $dG(y)$ は Stieltjes 積分であること．これを確率密度函数だけ考えて，$f(x)dx$, $g(y)dy$ としたのでは定理は不成立．実際，大変多くの場合に，最適の手は確率密度函数では表わされない．もう1つの注意は，payoff 函数 $K(x, y)$ が連続であることである．不連続であると定理が成立しないことがある．

4. games of survival

多段 game は単なる一段 game の拡張としてばかりでなく，

それ自体が基本的と考えられる．一段 game は，dynamic である多段 game の静態的な1こまである．この節では，多段 game の一例として survival game につき説明する．

survival game は，古典的な破産の問題の1つの拡張といえよう．2人の賭博者，I, II がいて，始めそれぞれ資本 r, $R-r$ をもっている．銅貨（必ずしも公平ではない）を投げて表が出たら，IIからIへ1，裏が出たら -1，を払う．これを何回も繰返して先に破産したものが負となる．この問題では手 (strategy) の考えは全く介入しない．銅貨の表の出る確率を p とすると，I，IIが負ける確率，ちょうど n 回で勝負がつく確率などが，確率論だけの計算から求まる．(たとえば Feller の本の第14章)．

ところが，これを次のように変えてみよう．player I，II がそれぞれ資本 r, $R-r$ をもって game を始めるのだが，毎回，銅貨を投げるという偶然試行をやるのではなく，1つの行例 game を行なうのだとする，もしもI，IIがそれぞれ手 $\alpha_i(i=1,\cdots\cdots,m)$, $\beta_j(j=1,\cdots\cdots,n)$ を出せば，IIから I へ a_{ij} 渡すとする．毎回この行列 game を続けて，先に相手を破産させた方が勝である．全 play 終了したときに payoff が勝てば，1，負ければ，0，引分け（どちらの player も永久に破産しないこともある．このときは引分けとする）も0とする．この game の特徴は

(i) 全 game が何段で終るかは，偶然 (chance) によってではなく，players の用いる手によって支配される．

(ii) payoff が全 play の終了時に生ずる．毎回の部分(行列)game のすんだあとではない．

であることに注意する．

このような特性をもった game を survival game という．次に簡単な survival game につき数学解析をやってみる．payoff が前記のようにきめてあるから，Iの所得の期待値は，Iが勝つ確率と同じ．例によって

$f(r)$……Iが資本 r のとき，最適政策を用いて逐に勝つ確率

とすると，$f(r)$ は次の函数方程式をみたす．

$$f(r)=\max_{p}\ \min_{q}\ \sum_{j=1}^{n}\sum_{i=1}^{m}\quad f(r+a_{ij})\ p_i q_j=\min_{q}\ \max_{p}\ [\cdots] \tag{4}$$

境界条件として

$$f(r)=0\ \ (r\leqq 0),\ \ =1\ (r\geqq R).$$

(4) 右辺は $(f(r+a_{ij})\mid i=1,\cdots,m;\ j=1,\cdots,n)$ を payoff 行列とする行列 game の値であるが，この game の最適の手を p^0, q^0 とすると，これは何れも r の函数である．r の函数としてきめられた p^0, q^0 が，求める最適政策であることはいうまでもない．

さて函数方程式 (4) を解こうとするのだが，先ず無意味な場合を除くために $A=(a_{ij})$ が，正の行，または負の列が存在しないようなものに限ってよい．何故ならば，もし A に正の行があれば，Iはその手 (mixed でなく pure である) ばかりを出せば，資産が増える一方で遂に勝つ

に決っている．もし A に負の列があれば，IIがその手ばかりを出せば，Iの資産は減る一方で遂にIが負けるに決っている．

そこで簡単のために

$$A=\begin{pmatrix} -1 & a \\ c & -b \end{pmatrix}, \qquad a,\ b,\ c \text{ は正整数}$$

で，r, R も正整数とすると，(4) は

$$f(r)=\begin{cases} 0, & r\leqq 0 \\ \mathrm{val}[(f(r+a_{ij}))], & r=1,2,\cdots,\ R-1 \\ 1, & r\geqq R \end{cases} \tag{5}$$

となる．行列 game $(f(r+a_{ij}))$ の値を val […] で表わしたのである．

定理 1. 函数方程式 (5) の解が一意に存在する．

証明 解の存在は容易である．

$$f_0(x)=\begin{cases} 0, & x\leqq R-1 \\ 1, & x\geqq R \end{cases}$$

$$f_{n+1}(x)=\begin{cases} 0, & x\leqq 0 \\ \mathrm{val}[(f_n(r+a_{ij}))], & 1\leqq x\leqq R-1 \quad ; \quad n=0,\ 1,\ 2,\cdots \\ 1, & x\geqq R \end{cases}$$

とおくと，$f_0(x)\leqq f_1(x)$，帰納法により，すべて，n, x に対して $f_n(x)\leqq f_{n+1}(x)$．もちろん $0\leqq f_n(x)\leqq 1$ だから，$f_n(x)$ はある函数 $f(x)$ に収束する．val 函数の連続性から，この $f(x)$ が (5) をみたすことは明らか．

次に解の一意性を示そう．かりに他の解 $g(x)$ があって，

$$\Delta\equiv\max_{1\leqq x\leqq R-1}|f(x)-g(x)|>0 \tag{*}$$

とせよ，今，$\Delta=|f(y)-g(y)|$，この y に対する f, g の optimal choices をそれぞれ $p(y)$, $q(y)$; $\bar{p}(y)$, $\bar{q}(y)$ とすると，

$$f(y)=\sum_{i,j} f(y+a_{ij})p_i(y)q_j(y),$$

$$g(y)=\sum_{i,j} g(y+a_{ij})\bar{p}_i(y)\bar{q}_j(y)$$

だから

$$\sum_{i,j}\{f(y+a_{ij})-g(y+a_{ij})\}\bar{p}_iq_j\leqq f(y)-g(y)\leqq\sum_{i,j}\{f(y+a_{ij})-g(y+a_{ij})\}p_i\bar{q}_j.$$

この両辺は $|\cdots|\leqq\Delta$ なのだから，少くともどちらか一方の等号が成立たねばならない．いま，左方の等号を成立とする：

$$\Delta=|f(y)-g(y)|=\left|\sum_{i,j}\{f(y+a_{ij})-g(y+a_{ij})\}\bar{p}_iq_j\right|.$$

$\bar{p}$, q は確率 vector であるから，$|f(y+a_{ij})-g(y+a_{ij})|<\Delta$ ならば $\bar{p}_iq_j=0$. 故に y を

始めから，(*) で max. に到達する最大整数にとっておけば，$a_{ij}>0$ ならば $\bar{p}_i q_j=0$. これは与えられた A に対して

$$\bar{p}_1 q_2=\bar{p}_2 q_1=0 \qquad \therefore\ q_1 \text{ or } q_2=0$$

となるが，q は行列 game

$$\begin{pmatrix} f(y-1) & f(y+a) \\ f(y+c) & f(y-b) \end{pmatrix}$$

の min-max strategy であったから，この証明の最後に示すように $f(x)$ が単調増大（非減少でない）なことにより，q_1 or $q_2=0$ ということはあり得ない．これで矛盾がでた．

最後に $f(x)$ の単調増大性：

$$0=f(0)<f(1)<f(2)<\cdots<f(R)=1 \tag{**}$$

の証明．$k\equiv\max\{\times\,[f(x)=0\},\ 1\leqq k\leqq R-1$ とおくと

$$f(k)=\mathrm{val}\left[\begin{pmatrix} f(k-1) & f(k+a) \\ f(k+c) & f(k-b) \end{pmatrix}\right]=\mathrm{val}\left[\begin{pmatrix} 0 & f(k+a) \\ f(k+c) & 0 \end{pmatrix}\right]>0.$$

これは矛盾であるから $k=0$. 故に $f(1)>0=f(0)$.

次に $f(x)$ の単調非減少性は明らかであることに注意する．それは $f_n(x)$ の定義から，$f_n(x)$ $(n=0,\ 1,\ 2,\cdots\cdots)$ が皆そうで，極限に行った $f(x)$ もそうだから，そこで，$f(1)>0, f(a+2)\geqq f(a+1)>0, f(c+2)\geqq f(c+1)>0,\ f(2-b)\geqq 0$. 故に

$$f(2)=\mathrm{val}\left[\begin{pmatrix} f(1) & f(2+a) \\ f(2+c) & f(2-b) \end{pmatrix}\right]\geqq\mathrm{val}\left[\begin{pmatrix} f(1) & f(1+a) \\ f(1+c) & 0 \end{pmatrix}\right]$$

$$>\mathrm{val}\left[\begin{pmatrix} 0 & f(1+a) \\ f(1+c) & 0 \end{pmatrix}\right]=f(1)$$

以下同様にして (**) 全部が証明される．こうして全証明が終った．

この特殊な問題に対するこの証明法は，しかし大変一般的なものである．この証明法は種々のこの種の函数方程式に使える．解の存在を示すのは易しかったが，その一意性を示すのは相当に難しく，その主要な部分は解の strict monotonicity に集中しているのである．

一般の方程式 (4) において，x が a_{ij} に比べて大きいときを考える．

$$f(x+a_{ij})\simeq f(x)+a_{ij}f'(x)$$

とかけるとすると，(4) は

$$f(x)\simeq\mathrm{val}\,[(f(x)+a_{ij}f'(x))].$$

もしも，

$$f'(x)>0$$

ならば（前定理で証明したように，離散型の問題に対して $f(x)$ の単調増大性があったから，連続型のときにも $f'(x)>0$ を想定するのは妥当である），上式右辺が更に

$$=f(x)+f'(x)\ \mathrm{val}\,[(a_{ij})]$$

となるから，結局

$$\mathrm{val}\,[(a_{ij})] \simeq 0 \tag{5}$$

を得る．この式の意味は次の通り：x が a_{ij} に比べて大変大きくて，全 play が終るまでに多数段階が残っているとすると，最適の手は単なる一段 game (a_{ij}) のそれと近似的に等しい（x に無関係である）．

a_{ij} を x に比べて小さくとるということは，本質的には，決定過程を連続型に転化しようとするのである．しかし，連続型の多段 game の数学解析は，一般に非常に面倒なことが知られている．

5. 非-O-和 2-人 game と拡張された min-max 定理

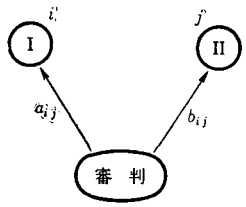
先ず非-O-和 2-人 game について説明しよう．§2 で述べた O-和 2-人 game では

$$b_{ij} = -a_{ij} \tag{1 再}$$

であるから，Ⅰが審判から得た a_{ij} は，ちょうどⅡが審判に渡したと同じものになっている．つまり，Ⅱから a_{ij} が審判を素通りしてⅠに渡されたことになり，審判は始めから存在しないものと考えてよい．ところが，もし (1) が成立しないとすると，すなわち行列記法で

$$A + B \neq 0 \tag{7}$$

（右辺は零行列）とすると，審判を省略して考えることはできなくなる．Ⅰ，Ⅱの収支は，一般には O-和のときのように正反対ではない．すなわち，$A+B$ が定数行列（要素が皆一定の数 c より成る行列）に等しい場合を除けば，Ⅰ，Ⅱの利害は複雑に相関する．



この場合著しいことは，game の均衡基準 (equilibrium criterion) を定めることが困難なことである．O-和 game の均衡基準は pAq で，Ⅰは p を動かして最大に，Ⅱは q を control して最小にしようと行動して，min-max 定理によってそこに均衡が生ずる．しかし，非-O-和 game では，事情はそのように単純ではない．Ⅰにとっては，Ⅱが A の所得 pAq を最小にしようとするのか，あるいは自分の所得 pBq を最大にしようとするのか，何れを意図しているのか不明である．Ⅱにとっても同様．このような事態における player の，何らかの意味で合理的な行動の基準をきめることは，易しくはない．

Nash (1950) は均衡の定義を拡張することによって難点を打開した．もしも

$$\left.\begin{aligned} p^*Aq^* &= \max_p\ pAq^* \\ p^*Bq^* &= \max_q\ p^*Bq \end{aligned}\right\} \tag{8}$$

のような手の組 $(p^*,\ q^*)$ が存在すれば，これをこの非-O-和 2-人 の均衡点 (equilibrium point) という．p^*, q^* は各人のいわば“きめ手”で，相手にそれを出されることを覚悟すれば，自分もそれを出さなければ必らず損，なのである．特に $B=-A$ とおくと $\min_q p^*Aq = p^*Aq^* =$

$\max_p pAq^*$ となるから，p^*, q^* は最適の手になる．均衡点は O-和 game の解（最良の手の組 (p^0, q^0) を解 (solution) という）の概念を拡張したものである．やはり

基本定理 非-O-和 2-人 有限 game には必らず均衡点が存在する (Nash, 1950)

が成立する．

本節では，Nash のものとはまた別の均衡基準を示してみよう．

前節の survival game で，始めの資本をそれぞれ x, y, 毎母の成分 game を，A, B できまる非-O-和 2-人 game に変える．

$f(x, y)$……I，IIがそれぞれ資本 x, y のとき，両者が共に最適政策を用いて，遂にIが勝つ確率とすると，

$$f(x, y) = \mathrm{val}\ [(f(x + a_{ij}, y+b_{ij}))] \tag{7}$$

が成立する．境界条件は

$$f(x, y) = 0 \quad (x \leqq 0 < y), \ = 1\ (y \leqq 0 < x), \ = \frac{1}{2}(x = y = 0).$$

最後の $x = y = 0$ の場合は引分けに相当する．値を $1/2$ としたのは単に便宜上のものである．

定理 2. すべての i, j に対して $a_{ij} + b_{ij} < 0$ ならば，

(7) に一意の解が存在する．

が成立する．証明は前節の定理のそれと殆んど同じ．

再び，$|a_{ij}| \ll x$, $|b_{ij}| \ll y$ と仮定すると

$$f(x + a_{ij}, y + b_{ij}) \simeq f(x, y) + a_{ij} f_x + b_{ij} f_y$$

とかけて，更に

$$f_{x,\ j}\ f_y > 0$$

と仮定すると，(7) は

$$\begin{aligned} f(x, y) &\simeq \mathrm{val}\ [(f(x, y) + a_{ij} f_x + b_{ij} f_y)] \\ &= f(x, y) + f_x\ \mathrm{val}\ [(a_{ij} + b_{ij} f_y/f_x)] \\ &= f(x, y) + f_y\ \mathrm{val}\ [(a_{ij} f_x/f_y + b_{ij})]. \end{aligned}$$

結局

$$\left.\begin{aligned} \mathrm{val}\ [(a_{ij} + b_{ij} f_y/f_x)] &\simeq 0 \\ \mathrm{val}\ [(a_{ij} f_x/f_y + b_{ij})] &\simeq 0 \end{aligned}\right\} \tag{8}$$

が出る．後に証明するが

定理 3. A, B は同じ大きさの行列で

$$\min_{i,j} b_{ij} \equiv b > 0 \tag{9}$$

とする．すると u に関する方程式

$$\mathrm{val}\ [A - uB] = 0 \tag{10}$$

は唯一つの根を有し，それは

$$u = \max_p \min_q \frac{pAq}{pBq} = \min_q \max_p \frac{pAq}{pBq} \tag{11}$$

で与えられる．

があるので，a_{ij}, b_{ij} は皆 <0 と仮定すると，(8) から

$$\left.\begin{aligned} -f_y/f_x &\simeq \max_q \min_p \frac{pAq}{pBq} = \min_p \max_q \\ -f_x/f_y &\simeq \min_q \max_p \frac{pBq}{pAq} = \max_p \min_q \end{aligned}\right\} \tag{12}$$

が得られる，一般に

$$\frac{1}{\max_q \min_p [\quad]} = \min_q \max_p \frac{1}{[\quad]}$$

であるから，(12) の二式は両立し得る．

定理の (11) 式によって，非-O-和 2-人有限 game の1つの均衡基準として

$$\frac{pAq}{pBq} \tag{13}$$

をとることが提案される．しかし，これを受け容れるかどうかは好みの問題であろう．両players が同一の基準函数を用いているかどうか先験的に明らかではないし，またもっと悪いことには，両者の効用尺度 (utility scale) が違うかも知れないような非-O-和 game では，かかる問題はいつも起りがちである．

定理3の証明　(10) が唯一つの根をもつことが証明できさえすれば，(11) を示すのはやさしい．実際，$A-uB$ の max-min strategy の1つ p^0 をとると

$$p^0(A-uB)q \geqq 0 \quad より \quad u \leqq \max_p \min_q \frac{pAq}{pBq}.$$

$A-uB$ の min-max strategy の1つ q^0 をとると

$$p(A-uB)q^0 \leqq 0 \quad より \quad u \geqq \min_q \max_p \frac{pAq}{pBq}.$$

かくて (11) が出た．

さて，(10) が一意に解けることを証明するには，逐次近似法を使うのが便利である．函数方程式系

$$\begin{aligned} u_0 &= \mathrm{val}\, A \\ u_{n+1} &= \mathrm{val}\,[A + u_n(I-B)], \qquad n \geqq 1 \end{aligned}$$

を考える．第2式で I としたものは要素が1ばかりである A, B と同大の行列である．

$$\begin{aligned} u_{n+1} &= \bar{p}(A + u_n(I-B))q \equiv T_n(\bar{p}, \bar{q}) \\ u_n &= \tilde{p}(A + u_{n-1}(I-B))\tilde{q} \equiv T_{n-1}(\tilde{p}, \tilde{q}) \end{aligned}$$

とおくと

$$(u_n - u_{n-1})\bar{p}(I-B)\bar{q} = T_n(\bar{p},\ \bar{q}) - T_{n-1}(\bar{p},\ \bar{q}) \leqq u_{n+1} - u_n$$
$$\leqq T_n(\bar{p},\ \bar{q}) - T_{n-1}(\bar{p},\ \bar{q}) = (u_n - u_{n-1})\bar{p}\ (I-B)\bar{q}$$

故に

$$|u_{n+1} - u_n| \leqq |u_n - u_{n-1}| \max\ (\bar{p}(I-B)\bar{q},\ \bar{p}(I-B)\bar{q}) \leqq |u_n - u_{n-1}|(1-b),$$

したがって $\sum(u_{n+1} - u_n)$ が幾何収束し，$\lim u_n$ が存在する．

一意なことをいうには，他の解を v とすると同様にして

$$|u - v| \leqq |u - v|\ (1-b)$$

が出るから　　$|u - v| = 0$　となる．

この定理で条件 (9) は不可決である．たとえば

$$A = \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

とすると

$$\mathrm{val}\,[A - uB] = \mathrm{val}\left[\begin{pmatrix} 0 & 2-u \\ 1-u & 0 \end{pmatrix}\right] = \frac{(1-u)(2-u)}{(1-u) + (2-u)}$$

となって，この函数の零点は 1 および 2.

6. 多段 game と情報の問題

ここで多段決定過程の定義をもう一度振返ってみよう．物理的位置の系 $D = \{p\}$ と，その上の self-mapping な変換の集合 $T = \{T_k\}$ が与えられている．初期位置 p^0 から出発して次々と

$$T_{k1}(p^0) = p^1,\ T_{k2}(p^1) = p^2, \cdots\cdots,\ T_{kn}(p^{n-1}) = p^n, \cdots\cdots$$

と，T_k を撰んで p を変化させてゆくとき，N 段階後の return $R(p^N)$ $(N \leqq \infty)$ を最大または最小にするような政策，すなわち T_k の撰択の列を求めること．

注意すべきことは，各段階において，その前段階で下した撰択（決定）の結果残された位置 (resulting position) を，decision maker が知っていることである．知っていてこそ始めて，この段階での適切な決定が下せるわけである．

多段 game でもこのことは同じである．各段階での decision maker は互いに対抗する2主体—players—である．両 player は各段階において，その直前の段階での成分 game がいかに戦われたか，またその結果残された状態を知っている．これが多段 game の構造の特徴で，もしこのことがなければ，それは本質に一段 game に帰する．

この章では，多段 game と情報の問題について興味ある一定理を述べよう．$N \times m$ 行列

$$\begin{pmatrix} a^{(1)} \\ a^{(2)} \\ \vdots \\ a^{(N)} \end{pmatrix}, \qquad (\text{各 } a^{(i)} \text{ は } m \text{ 次元の行 vector})$$

のある与えられた集合 A が，条件

(i) すべての $\begin{pmatrix} a^{(1)} \\ a^{(2)} \\ \vdots \\ a^{(N)} \end{pmatrix} \in A$ に対して $\sum_1^N a^{(i)} = a.$

(ii) $(1, 2, \cdots, N)$ のすべての順列 $(\pi_1, \cdots, \pi_N)$ に対し,

$$\begin{pmatrix} a^{(1)} \\ a^{(2)} \\ \vdots \\ a^{(N)} \end{pmatrix} \in A \quad \text{ならば} \quad \begin{pmatrix} a^{(\pi_1)} \\ a^{(\pi_2)} \\ \vdots \\ a^{(\pi_N)} \end{pmatrix} \in A.$$

をみたすとする．同様に，$n \times N$ 行列

$$[b^{(1)}\ b^{(2)}\ \cdots\ b^{(N)}], \qquad (b^{(i)} \text{ は } n \text{ 次元の列 vector})$$

のある集合 B も (i) (ii) をみたす (今度は a の代りに b) とする．N 段 game が次のように行われる：

段	I	II
第　段		$[b^{(1)}\ b^{(2)} \cdots\cdots b^{(N)}] \in B$ を撰ぶ
1	$a^{(1)}$ を撰ぶ	$b^{(1)}$ を I に知らせる
2	$a^{(2)}$ を撰ぶ	$b^{(2)}$ を I に知らせる
⋮	⋮	⋮
N	$a^{(N)}$ を撰ぶ	$b^{(N)}$ を I に知らせる

ただし，第 N 段で I は

$$\begin{pmatrix} a^{(1)} \\ \vdots \\ a^{(N)} \end{pmatrix} \in A$$

になっていなければならぬ．payoff は II から I へ

$$\sum_{i=1}^{N} a^{(i)} M b^{(i)},$$

M は与えられた $m \times n$ 行列である．

定理 4. この多段 game の値は

$$(1/N)\, a M b$$

である．I の最適の手は，任意の $\begin{bmatrix} a^{(1)} \\ \vdots \\ a^{N} \end{bmatrix}$ に対するすべての $\begin{bmatrix} a^{(\pi_1)} \\ \vdots \\ a^{(\pi_N)} \end{bmatrix}$ を等確率 $1/N!$ でとる．II のそれも同様. (Gale, 1957)

証明 この II の最適の手を τ^0 とかく，I の任意の pure strategy σ に対して，I の所得の期待値は

$$E(\sigma,\ \tau_0) = (1/N)\, a M b$$

であること帰納法で示そう．

$N=1$ のとき成立は明らか，$N-1$ のとき成立とせよ．Ⅰの第1段の手が $a^{(1)}$ とすると，

$$E(\sigma,\ \tau^0)=\frac{1}{N}\sum_{k=1}^{N}\left\{a^{(1)}Mb^{(k)}+\frac{(a-a^{(1)})M(b-b^{(k)})}{N-1}\right\}=\frac{1}{N}aMb$$

となるから，N でも成立．右辺の因子 $1/N$ は，τ^0 によりⅡの手の第一成分が $b^{(k)}$ となる確率である．

全く同様にして，Ⅰの最適の手を σ^0 とかくと，Ⅱの任意の pure Strategy τ に対して

$$E(\sigma^0,\ \tau)=(1/N)\,aMb$$

が示される．

この多段 game は“当てっこゲーム”(guess game) の型である．Ⅱ (nature のこともある) は第0段だけで決定をなし，第 $1\sim N$ 段ではⅠだけが決定をくだす．段階 i では，ⅠはⅡからそれまでに $b^{(1)},\cdots,b^{(i-1)}$ について知らされているから，条件 (i) により，$b^{(i)}$ についての部分的情報を既に与えられているわけである．しかしこの定理は，Ⅰの最適政策はこれらの情報を全然利用しないことを示す!.

(注意： この事実は奇妙ではない．game の理論では，両 player が共に rational で最適の手をとるものと予想する．上述の場合でも，Ⅱが彼の最適政策をとるから，Ⅰの最適政策が情報を利用しないことになるのである．Ⅱが rational でなくて彼の最適政策をとらなかったら，多分Ⅰは情報を利用することによって $(1/N)aMb$ より多くを得られるであろう，このときⅠはⅡのバカにつけこむのである．)

7. 例 題 と 解 説

1. 2人の player Ⅰ，Ⅱ がいて，銅貨の片面をみせ合う (matching coin)．ルールは

i) Ⅰ，Ⅱ が共に表—表をみせたら，共に1を失う．

ii) Ⅰ，Ⅱ 共に裏—裏をみせたら，ⅡからⅠへ1やる，

iii) その他の場合は，ⅠからⅡへ1やる，

最初Ⅰが m，Ⅱが n (共に正整数) 持っていて，この見せ合いを幾回も続けて先に相手を破産させた方が勝ちとする．

Ⅰ，Ⅱが表を見せる確率をそれぞれ $x,\ y$ とすると

$q_1=xy$ ……表—表になる確率

$q_2=x(1-y)+y(1-x)$ ……表—裏，または裏—表になる確率

$q_3=(1-x)(1-y)$ ……裏—裏になる確率

である．仮定：

“毎回同じ手を使う”

(すなわち，$m,\ n$ に無関係に) とする．このときⅡが負けない (先にⅠを破産させるか，またはⅠと同時に破産する．Ⅱにとって勝つか引分け) 確率を $p(m,\ n;\ x,\ y)$ とすると

$$p(m,\ n;\ x,\ y)=q_1p(m-1,\ n-1;\ x,\ y)$$
$$+q_2p(m-1;\ x,\ y)+q_3p(m+1,\ n-1;\ x,\ y),\quad m,\ n\geqq 1.$$

境界条件は，すべての $0 \leqq x,\ y \leqq 1$ に対し

$$p\,(m,\ 0\,;\ x,\ y) = 0, \qquad m \geqq 1,$$

$$p\,(0,\ n\,;\ x,\ y) = 1, \qquad n \geqq 0.$$

1) たとえば

$$p(2,\ 1\,;\ x,\ y) = \frac{(q_1 + q_2)\,q_2}{1 - q_2 q_3} = \frac{(x + y - xy)\,(x + y - 2xy)}{1 - (1 - x)\,(1 - y)\,(x + y - 2xy)}$$

となる．

$$\min_{0 \leqq x \leqq 1}\ \max_{0 \leqq y \leqq 1} p(2,\ 1\,;\ x,\ y) \neq \max_{0 \leqq y \leqq 1}\ \min_{0 \leqq x \leqq 1} p(2,\ 1\,;\ x,\ y)$$

であることを示せ．

2) 連続 game の基本定理 (§3) から

$$\max_{G \in D}\ \min_{F \in D} \int_0^1 \int_0^1 p(2,\ 1\,;\ x,\ y)\,dF(x)\,dG(y) = \min_{F \in D}\ \max_{G \in D}\ [\ \cdots\ ]$$

であるが，このときの両者の最適の手が

1°. II に対しては

$$p(2,\ 1\,;\ 0,\ y^0) = p(2,\ 1\,;\ 1,\ y^0)$$

を満足する y^0 が，最適な pure な手になる．

2°. I に対しては

$$\max_y \{\, a p(2,\ 1\,;\ 0,\ y) + (1 - a) p\,(2,\ 1\,;\ 1,\ y) \}$$

$$= a p(2,\ 1\,;\ 0,\ y^0) + (1 - a)\,p\,(2,\ 1\,;\ 1,\ y^\circ)$$

を満足する a をとると

$$F^0\,(x) = \begin{cases} 0, & x = 0 \\ 1 - a, & 0 < x < 1 \\ 1, & x = 1 \end{cases}$$

(すなわち，それぞれ a, $1 - a$ の割合で全部表，全部裏を見せる) が，最適な混合手になり，

且つ，game の値が，方程式

$$y = \frac{1 - y}{1 - y + y^2}$$

の唯一の実根 $y_* \simeq 0.4302$ に等しい．これらのことを示せ．

3) 仮定 (*) を取去り，

$p(m,\ n)$……I，II がそれぞれを m, n もっているとき，最適政策を用いたときの，II が負けない確率

とおけば，今度は本当の dynamic programming の問題になって，方程式は

$$p(m,\ n) = \max_G \min_F \int_0^1 \int_0^1 \{ q_1 p(m-1,\ n-1) + q_2 p(m-1,\ n+1)$$

$$+ q_3 p\ (m+1,\ n-1)\}\ dF(x)\ dG(y)$$

$$= \min_F \max_G \left(\cdots\cdots \right) \qquad m,\ n \geqq 1.$$

境界条件は，前と同じく

$$p(m,\ 0) = 0, \qquad m \geqq 1,$$

$$p(0,\ n) = 1, \qquad n \geqq 0.$$

$m = 2,\ n = 1$ の場合にこれを解いてみよ．

(Bellman-Blackwell, 1949)

2.

$$A = \begin{pmatrix} 2 & -1 \\ -2 & 1 \end{pmatrix}$$

$$R = 4$$

である survival game について

$f(k)$……Ⅰの財産が k であるとき，最適政策を用いたときのⅠの survival の確率とすると

$$f\ (1) = \frac{f(2)\,f(3)}{f(2) + f(3)},$$

$$f\ (2) = \frac{f(3)}{1 + f(3) - f(1)},$$

$$f\ (3) = \frac{1 - f(1)\,f(2)}{2 - f(1) - f(2)},$$

したがって

$$f(1) = 1 - \frac{\sqrt{2}}{2} \simeq 0.293,\ f(2) = \frac{1}{2},\ f(3) = \frac{\sqrt{2}}{2} \simeq 0.707,$$

最適の手は

Ⅰの財産	Ⅰ		Ⅱ	
$r = 1$	$(0.414\,\alpha_1,$	$0.586\,\alpha_2)$	$(0,414\,\beta_1,$	$0.586\,\beta_2)$
2	$(0.5\ \ \alpha_1,$	$0.5\ \ \alpha_2)$	$(0.293\,\beta_1,$	$0.707\,\beta_2)$
3	$(0,586\,\alpha_1,$	$0.414\,\alpha_2)$	$(0.414\,\beta_1,$	$0.586\,\beta_2)$

となることを示せ，

(Hausner, 1952)

(解)　方程式は

$$f(k)=\begin{cases}0, & k\leqq 0,\\ \mathrm{val}\left[\begin{pmatrix} f(k+2) & f(k-1)\\ f(k-2) & f(k+1)\end{pmatrix}\right], & k=1,\ 2,\ 3,\\ 1, & k\geqq 4\end{cases}$$

である．したがって

$$f(1)=\mathrm{val}\begin{pmatrix} f(3) & 0\\ 0 & f(2)\end{pmatrix}=\frac{f(2)f(3)}{f(2)+f(3)}$$

$$f(2)=\mathrm{val}\begin{pmatrix} 1 & f(1)\\ 0 & f(3)\end{pmatrix}=\frac{f(3)}{1+f(3)-f(1)}$$

$$f(3)=\mathrm{val}\begin{pmatrix} 0 & f(2)\\ f(1) & 1\end{pmatrix}=\frac{1-f(1)f(2)}{2-f(1)-f(2)},$$

これらから $f(1)$, $f(2)$, $f(3)$ を解けばよい．

3. $$A=\begin{bmatrix} a & -1\\ -1 & b\end{bmatrix},\qquad (a,\ b \text{ は正整数})$$

$$R=n$$

である suevival game について

$v_n(k)$……Ⅰの財産が k であるとき，最適政策を用いたときのⅠの survival の確率とする．次の諸事項を示せ．

1) $$v_n(1)=\frac{v_n(1+a)\ v_n(1+b)}{v_n(1+a)+v_n(1+b)}$$

$$p_n(1)=\frac{v_n(1+b)}{v_n(1+a)+v_n(1+b)}=\frac{v_n(1)}{v_n(1+a)}$$

である．$k=1$ のときのⅠの最適の手を $(p_n(1)\alpha_1,\ (1-p_n(1))\alpha_2)$ とした．

2) 一般に survival game について，漸化式

$$v_{n+1}(k+1)=v_n(k)+(1-v_n(k))v_{n+1}(1),\qquad 1\leqq k<R$$

が成立する．

3) したがって

$$v_{n+1}(1)=\frac{\sqrt{v_n(a)v_n(b)}}{1+\sqrt{v_n(a)v_n(b)}}$$

となる．

(解) 1)

$$v_n(k)=\begin{cases}0, & k\leqq 0\\ \mathrm{val}\left[\begin{pmatrix} v_n(k+a) & v_n(k-1)\\ v_n(k-1) & v_n(k+b)\end{pmatrix}\right], & k=1,\cdots,n-1\\ 1, & k\geqq n\end{cases}$$

故に

$$v_n(1) = \mathrm{val}\begin{pmatrix} v_n(1+a) & 0 \\ 0 & v_n(1+b) \end{pmatrix} = \frac{v_n(1+a)\ v_n(1+b)}{v_n(1+a) + v_n(1+b)}$$

2) $v_{n+1}(k+1)$……両 player の財産の状態が $(k+1,\ n-k)$ のときから始めて，遂にⅡが破産する確率

$v_n(k)$……両 player の財産の状態が $(k,\ n-k)$ のときから始めて，遂にⅡが破産する確率

である．更に，$R=n+1$ のときの survival game でⅠが勝ち残ったときに，Ⅱが破産するまでに財産の状態 $(1,\ n)$ を通ったか，通らなかったかの何れかである．後の場合は，Ⅰの資本のうち1だけは全然 game に介入しなかったことになるから，この確率は $v_n(k)$ に等しい．前者の場合の確率が $(1-v_n(k))v_{n+1}(1)$ である．

3)
$$v_{n+1}(1) = \frac{v_{n+1}(1+a)v_{n+1}(1+b)}{v_{n+1}(1+a) + v_{n+1}(1+b)}$$

$$v_{n+1}(1+a) = v_n(a) + (1-v_n(a))v_{n+1}(1)$$
$$v_{n+1}(1+b) = v_n(b) + (1-v_n(b))v_{n+1}(1)$$

より $v_{n+1}(1)$ を解く．$v_{n+1}(1)$，$v_n(a)$，$v_n(b)$ をそれぞれ，y，α，β とおくと

$$y = \frac{\alpha\beta + ((1-\alpha)\beta + \alpha(1-\beta))y + (1-\alpha)(1-\beta)y^2}{\alpha+\beta+(2-\alpha-\beta)y}$$

整頓して

$$(1-\alpha\beta)y^2 + 2\alpha\beta y - \alpha\beta = 0$$

$$\therefore\ y = \frac{-\alpha\beta \pm \sqrt{\alpha\beta}}{1-\alpha\beta}$$

負根はすてると

$$y = \frac{\sqrt{\alpha\beta}}{1+\sqrt{\alpha\beta}}.$$

4. player Ⅰ，Ⅱ が資源をそれぞれ x，y だけもっている．両者がそれらを n 個の部分にわけて

$$x = \sum_1^n x_i, \qquad x_i \geqq 0 \tag{*}$$

$$y = \sum^n y_j, \qquad y_j \geqq 0 \tag{**}$$

としたとき，Ⅱ からⅠへの payoff が

$$M(x_1, \cdots\cdots, x_n;\ y_1, \cdots\cdots,\ y_n) = \sum_1^n c_j \max(x_i - y_j,\ 0)$$

とする．もちろんこれを，Ⅰは最大にⅡは最小にしようとする．

$f_n(x,\ y)$……両者が最適の手を用いたときのIの所得の期待値とする．

この game が §6 で説明したように $n-1$ 段 game ならば，関係式

$$f_n(x,\ y)=\max_{F\in D}\ \min_{G\in D}\int_0^x\int_0^y\{c_n\ \max(x_n-y_n,\ 0)$$

$$+f_{n-1}(x-x_n,\ y-y_n)\}dF(x_n)\ dF(y_n)$$

$$=\min_G\ \max_F\ [\ \cdots\ ] \qquad (***)$$

が成立する．単なる1段 game ならば，両者の手はそれぞれ (*)，(**) の上での $n-1$ 次元確率分布で表わされて

$$f_n(x,\ y)=\max_F\ \min_G\int\cdots\int M(x_1,\cdots,x_n;\ y_1,\cdots,\ y_n)$$

$$dF(x_1,\cdots,x_n)\,dG(y_1,\cdots,\ y_n)$$

$$=\min_G\ \max_F\ [\ \cdots\ ]$$

が成立する．(***)のような漸化式は成立しない．　　(連続型 Blotto game)

5. 表の出る確率 p が未知である銅貨があって，それを無限回投げるとする．毎回，前回までに出た結果を皆知って，今度出る面をあてる．当りの数の期待値を最大にするにはどうすればよいか．ただし，この数を収束させるために，第2段以後には割引率 $a(0<a<1)$ をかけて数えるものとする．また未知である p の事前確率分布を $dF(p)$ とせよ．

$f(m,\ n)$……m回表，n 回裏出たことを知ったときに，以後最適政策を用いて得られる期待値

とすると，これの満足する関係式を求めよ．p の事前分布を想定せずに，nature を対抗者として game 論的に扱うとどうなるか．

[解]

$$f(m,\ n)=\max\left\{\begin{array}{l}\int_0^1\{p(1+af(m+1,\ n))+(1-p)af(m,\ n+1)\}dF_{m,n}(p)\\ \int_0^1\{(1-p)(1+af(m,\ n+1))+paf(m+1,\ n)\}dF_{m,n}(p)\end{array}\right\},$$

ただし $dF_{m,n}(p)$ は事後確率で

$$dF_{m,\ n}(p)=\frac{p^m(1-p)^n dF(p)}{\int_0^1 p^m(1-p)^n dF(p)}.$$

あるいは

$$b(m,\ n)=\int p\,dF_{m,n}(p)$$

とおいて

$$f(m,\ n)=\max\left[\begin{array}{l}(1+af(m+1,\ n))b(m,\ n)+(1-b(m,\ n))af(m,\ n+1)\\ (1-b(m,\ n))(1+af(m,\ n+1))+b(m,\ n)af(m+1,\ n)\end{array}\right].$$

6. 定理4を利用して次の問題をとけ，

1) N枚のカードに数 $\beta_l(l=1,\ 2,\cdots,N)$ をかいて伏せてある．player が財産1だけもって

いて，α_1だけどれかのカードに賭ける．開いてみて β_i だったら $\alpha_1\beta_i$ もろう．残り $1-\alpha_1$ の一部をまた残り $(N-1)$ 枚のどれかに賭ける．このようにして N 回の所得の計を最大にせよ．

2) N個の箱のうち，D個には "dollar" が入っている．C 個をえらんで中の dollar の計を最大にせよ．利用してよい情報として，1 個づつ中を開けてみてもよい．ただし，その内容物はもしあっても数えないし，この箱はえらべない．　(Gale 1957)

[解] 1) payer II を仮想して，これがカードを一枚ずつ手渡すと考えよ．a, b, M は皆，数でそれぞれ 1, $\sum_1^N \beta_i$, 1.

$$A=\left\{\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_N\end{pmatrix}\,\middle|\,\sum_1^N \alpha_i=1,\ \alpha_i\geqq 0\right\}$$

$$B=\{(\beta_{\pi 1},\cdots\cdots,\beta_{\pi N})\}$$

値は　$(1/N)\sum_1^N \beta_i$.

2) player II を仮想して，これが箱を1個づつ手渡すと考える．

$$b^{(i)}=\begin{pmatrix}0\\ \vdots\\ 1\\ 0\end{pmatrix}\quad (j$$

の意味は

i 回目に手渡した箱が，$1\leqq j\leqq D$ ならば "dollar" であった．$D+1\leqq j\leqq N$ ならば空であた．

とする．次に

$$a^{(1)}=(0,\cdots\cdots,0,\ \overset{i}{1},\ 0)$$

は1回目に手渡された箱を，$1\leqq i\leqq C$ ならばえらぶ，$C+1\leqq i\leqq N$ ならば開けてみる

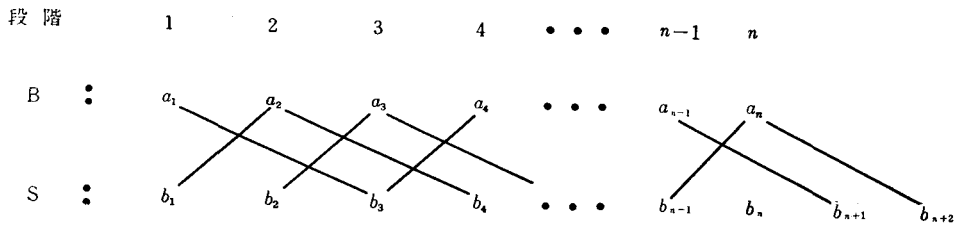
という手とする．

$$a=(1,\ 1,\cdots\cdots,\ 1).\qquad b=\begin{bmatrix}1\\ \vdots\\ 1\end{bmatrix}$$

$$M=\begin{array}{c|c|c|}
 & \rightarrow D\leftarrow & \\ \hline
\uparrow & & \\
C & \text{皆 } 1 & 0 \\
\downarrow & & \\ \hline
 & 0 & 0 \\ \hline
\end{array}$$

7. 船 S が1単位づつ右または左に動く．爆撃機 B がそれを fire する． fire してから2単位時間後に弾は海面にとどく．当れば，1, 外れれば0とする．弾は1発しかない．両者の最適政策を求めよ．

B(player I) の行動 (action) を a_1, a_2, a_3, ……, S(player II) の行動を b_1, b_2, b_3, ……とする.



上図は，たとえば第4段目で，B が a_4 の決定をくだすときには a_1, a_2, a_3 および b_1, b_2, b_3の知識がある．しかし S が b_4 の決定をくだすに当っては，b_1, b_2, b_3 および a_1, a_2 の知識しか利用できないことを示す．また，B の a_1, S の b_1, b_2 は何れも，相手の過去の行動について情報なしにとらねばならない．

Bが第 n 段階までに fire せねばならぬとする．この n 段 game の値を g_n としよう．Sの第1段の決定が，もしも

"確率 x で左へ"

ときまっているときの game の値を $g_n(x)$ とすると

$$g_n = \min_{0 \leqq x \leqq 1} g_n(x)$$

は明らか，$g_n(x)$ のみたす関係式を求めてみる．

fire してから弾がとどくまでの時間2の間に，Sの変位は左2，0，右2の3通りしかないから，Bはこの中のどれかにねらって fire するか，または fire しないかの何れかである．Sの第2段の決定を

"前段で左ゆきだったら確率 c で左へ,

前段で右ゆきだったら確率 d で左へ."

とすると，B の第1段で左，2，0，右2をねらって fire するとき，命中の確率がそれぞれ

$$xc, \quad x(1-c) + (1-x)d, \qquad (1-x)(1-d)$$

である．それで次の関係式が得られる．

$$g_n(x) = \min_{0 \leqq c,\ d \leqq 1} \max \begin{cases} xc \\ x(1-c) + (1-x)d \\ (1-x)(1-d) \\ x_n g_{-1}(c) + (1-x) g_{n-1}(d) \end{cases} \qquad n \geqq 2.$$

右辺の一番下の式は，第1段でBが fire しなかったときに対応する．$g_1(x)$ は第1段で必らず fire しなければならない (fire しなければ必らず0だから，した方がよいに決っている) 場合だから

$$g_1(x) = \min_{0 \leqq c,\ d \leqq 1} \max \begin{cases} xc \\ x(1-c) + (1-x)d \\ (1-x)(1-d). \end{cases}$$

これらが求める関係式である．　(Karlin, 1957)