不確実性の下での最適危機管理について

岩本, 誠一
九州大学経済学部 : 教授

https://doi.org/10.15017/4494376

出版情報：經濟學研究. 61 (5/6), pp.1-18, 1996-03-30. 九州大学経済学会

不確実性の下での最適危機管理について

岩　本　誠　一

1．はじめに

社会システムや組織においては「安全」と「危険」は表裏一体をなしているが，最近の日本では危機意識の欠如がいろいろな型で叫ばれているようである。特に，国の内外から「国家としての日本および個人としての日本人は政治的にも自然災害に対しても危機管理意識が薄い」と指摘されるようになった。国内では予期せぬ火山噴火や大規模地震による災害が多発している。また銃器盗難，毒ガス攻撃などかつてなかった新しい形の社会問題も発生している。本論文では，このような問題を多期間にわたる不確実性の下での最適な危機管理という問題としてとらえ，多段階意志決定過程として考える。すなわち，数学的には動的確率制御過程の枠組みの中で最小型関数の期待値最大化を思考する。

以下では本質的なことを簡略に表現するために 2 期間の不確実性の下でのミニマックス化を行う。有限な $n(\geq 2)$ 期間問題についても同様に取り扱えるからである。いわゆるミニマックス思考とは将来にわたっての考えられる最悪の危険をできるだけ軽減して良い方向に導こうとすることである。一般に，危険度(リスク)，危機の度合や後悔の程度といった諸量は非加法的である。また，心理学においても精神的なショックの程度など非加法的な量が考えられる。これらの量の複数個の総体としては意志決定者にとっては単に足し合わせて総合計をだすのではなくて，それらの値のうちの最大の値ないしは最小の値で評価すべきであろう。以下では，値 0 が最も危機的な状態にあり，値 1 が最も平和的な状態にあることを示すものとする。ここでは多期間にわたって不確実性が存在する場合を対象にしているので，ランダムに変動する危険に対して最悪（最小）の危険を評価値と考え，この期待値をできるだけよくすることを考える。具体的にはパラメータ付きマルコフ推移システムにおいて各段評価の（危機の度合いを表す）値の<u>最悪値</u>の期待値を最大化するという意味で合理的期待形成理論を取り扱うことになる。すなわち，**最小型関数の期待値**最大化を行う。

他方，国民総生産，投入資金などの量に対してはその総合計には極めて妥当性がある。一般に，経済学におけるないしは経済的な量は実質的にも加法的である。1995年ノーベル経済学賞受賞者 R. Lucas 等による合理的期待形成理論では各段評価の<u>総和</u>の期待値および<u>現在割引き総価値</u>の期待値を最大化している[12],[19]。これらはマルコフ決定過程[16]ないしは確率的動的計画法[6]としてとらえられ，**加法型関数の期待値**の最適化の範疇に入る。

2．問題と定式化

いま，ある2段システムが，初期状態 $x_0 \in X$ から出発して制御マルコフ推移法則 $p(x_1|x_0, u_0)$, $p(x_2|x_1, u_1)$ に従って確率ツリーを構成しながら移り，ある確率で $x_2 \in X$ になり，そこで終了するとする。ただし，$u_0 \in U$ は状態 x_0 を観察して取る決定（行動ともいう）であり，$u_1 \in U$ は次の状態 x_1 に依存して取る決定である。このとき，第1段では決定 u_0 に依存した評価（リスク）$r_0(u_0)$ が課され，第2段では u_1 に関係した評価 $r_1(u_1)$ が課され，最終の第2段終了時点には終端状態 x_2 に依存したゴール（目標）評価 $r_G(x_2)$ が課されるとする。このとき，システム全体としてはこれら三つの評価値の最小値（意志決定者にとっての最悪値）$r_0(u_0) \wedge r_1(u_1) \wedge r_G(x_2)$ が総合評価として下されるものとする。すなわち，システムとしての危機の度合いを各段でのリスクのうちで最も悪い値と考える。不確実な状況の下でこの総合評価値（最も悪い評価値）を最大にするように行動するには，意志決定者が各段での状態に応じてどのように決定を取っていけばよいかが問題である。これを数学的に記述すると，次の期待値最大化問題になる：

$$\begin{array}{lll} \text{Maximize} & E[r_0(u_0 \wedge r_1(u_1) \wedge r_G(x_2)] & \\ \text{subject to} & \text{(i)} \quad x_{n+1} \sim p(\cdot|x_n, u_n) \qquad n=0, 1 & \qquad (1) \\ & \text{(ii)} \quad u_0 \in U,\ u_1 \in U & \end{array}$$

ただし E は，初期状態 x_0，マルコフ推移確率 $p(y|x, u)$ および政策 $\pi=\{\pi_0, \pi_1\}$ から履歴の直積空間 $H=X \times U \times X \times U \times X$ 上に唯一定まる確率測度 $P^{\pi}_{x_0}$ による**期待値作用素**である。この意味で E には上下に添字を付けて $E^{\pi}_{x_0}$ であらわすべきだが，以下簡単のために単に E で表して用いる。したがって，この最大化問題の評価関数の値は

$$\begin{aligned} & E[r_0(u_0) \wedge r_1(u_1) \wedge r_G(x_2)] \\ &= \sum_{(x_1,x_2) \in X \times X} \{[r_0(u_0) \wedge r_1(u_1) \wedge r_G(x_2)] p(x_1|x_0, u_0) p(x_2|x_1, u_1)\} \qquad (2) \end{aligned}$$

である。ここに次のことに注意すべきである。**政策** π を構成する成分 $\pi_0, \pi_1 : X \to U$ をそれぞれ**第1決定関数**，**第2決定関数**と呼ぶが，式(2)中の u_0, u_1 はそれぞれ決定関数 π_0, π_1 を通して

$$u_0 = \pi_0(x_0), \quad u_1 = \pi_1(x_1) \qquad (3)$$

で定まっている([4])。

3．埋め込みと再帰式

さて，問題(1)を動的計画法で解くことを考えよう。一般に動的計画法では，与問題をそれ自身を含む適度の大きさの問題群に埋め込み再帰式を導いて，それを解くことになる。そのために，ここでは不変埋没原理を用いる([3], [13])。すなわち，最小型評価系の直前に新しくパラメータ $\lambda \in [0, 1]$ を導入して，部分問題群

$$v^2(x_2; \lambda) = \lambda \wedge r_G(x_2)$$

$$v^1(x_1;\lambda)=\underset{\pi_1}{\mathrm{Max}}\sum_{x_2\in X}[\lambda\wedge r_1(u_1)\wedge r_G(x_2)]p(x_2|x_1,\ u_1) \tag{4}$$

$$v^0(x_0;\lambda)=\underset{\pi_0,\pi_1}{\mathrm{Max}}\sum_{(x_1,x_2)\in X\times X}\{[\lambda\wedge r_0(u_0)\wedge r_1(u_1)\wedge r_G(x_2)] \times p(x_1|x_0,\ u_0)p(x_2|x_1,\ u_1)\} \tag{5}$$

を定義する。上式(4)，(5)における最大化は1次元拡大された状態空間上での決定関数 $\pi_0, \pi_1 : X\times[0,1]\to U$ 全体にわたるものとする。式(4)においては $u_1=\pi_1(x_1;\lambda)$ であるから，$v^1(x_1;\lambda)$ における最大化は第2決定関数 π_1 を動かす。また，式(5)では $u_0=\pi_0(x_0;\lambda)$，$u_1=\pi_1(x_1;\lambda_1)$ である。ここに $\lambda_1=\lambda\wedge r_0(u_0)$。$v^0(x_0;\lambda)$ での最大化は政策 $\pi=\{\pi_0,\ \pi_1\}$ 全体で取る。

この埋め込みによって(5)に $\lambda=1$ を代入した問題は原問題(1)になっている。(厳密には，§6 拡大過程上の終端評価および§7 定式化と最適性の再検討を参照。)したがって，$v^0(x_0;1)$ が求める最大値である。

このとき，$X\times[0,1]$ 上の ((x,λ) の) 2変数最大値関数列 $\{v^0,\ v^1,\ v^2\}$ 間には次の再帰式が成り立つ[1]：

定理 1

$$\begin{aligned} v^2(x;\lambda)&=\lambda\wedge r_G(x)\\ v^1(x;\lambda)&=\underset{u\in U}{\mathrm{Max}}\sum_{y\in X}v^2(y;\lambda\wedge r_1(u))p(y|x,\ u)\\ v^0(x;\lambda)&=\underset{u\in U}{\mathrm{Max}}\sum_{y\in X}v^1(y;\lambda\wedge r_0(u))p(y|x,\ u). \end{aligned} \tag{6}$$

4．最適解の構成

次に，原問題(1)の最適解を考えよう。一般に最適化問題の最適解は最適値と最適（値を与える）点の対であるが，動的計画法による最適化問題の最適解は(i)最適値関数列，(ii)最適政策，(iii)最適行動の3点で与えられる。この3点を用いて所与の最適化問題の最適値と最適点が求められる。ここでは(i)**最適値関数列**とは2変数関数列 $\{v^0,\ v^1,\ v^2\}$ である。(ii)**最適政策** π^* とは，式(6)において最大値に到達する u の値をそれぞれ $\pi_0^*(x;\lambda)$，$\pi_1^*(x;\lambda)$ で表すことによって得られる最適決定関数の列 $\{\pi_0^*,\ \pi_1^*\}$ である。(iii)**行動**とは一般に状態と決定から成る交互列 $(x_0;\lambda_0)\to u_0\to(x_1;\lambda_1)\to u_1\to(x_2;\lambda_2)$ であり，特に**最適行動**とは最適値を与える行動をいう。従って，与問題の最適値は最後に求められる(最初の)最適値関数 $v^0(x;\lambda)$ に初期状態 $(x_0;\lambda_0)$ を代入した $v^0(x_0;\lambda_0)$ で与えられる。また，最適点は初期状態 $(x_0;\lambda_0)$ からの最適行動で与えられる([7],[8])。

それでは，このことを用いて，(1)の最適解を求めよう。まず，再帰式(6)を $v^2(x;\lambda)\to v^1(x;\lambda)\to v^0(x;\lambda)$ の順に解いて，最適値関数列 $\{v^0,\ v^1,\ v^2\}$ が得られる。同時に，最大値を与える u を求めて最適政策 $\pi^*=\{\pi_0^*,\ \pi_1^*\}$ が得られる。従って，初期状態 $(x_0;1)$ からの最適行動

$$(x_0;1)\to u_0^*\to(X_1^*;\lambda_1)\to U_1^*\to(X_2^*;\ \Lambda_2) \tag{7}$$

1) 付録2で証明を与える。

が，最適政策 π^* と状態推移法則 $p(y|x, u)$ によって拡大履歴空間

$$\tilde{H}=(X\times[0, 1])\times U\times(X\times[0, 1])\times U\times(X\times[0, 1]) \tag{8}$$

上に確率的に定まる。ただし

$$\begin{aligned} &u_0^*=\pi_0^*(x_0; 1), \quad X_1^*\sim p(\cdot|x_0, u_0^*), \quad \lambda_1=1\wedge r_0(u_0^*)=r_0(u_0^*) \\ &U_1^*=\pi_1^*(X_1; \lambda_1), \quad X_2^*\sim p(\cdot|X_1, U_1^*), \quad \Lambda_2=\lambda_1\wedge r_1(U_1^*). \end{aligned} \tag{9}$$

ここに，小文字は確定数，大文字は確率変数を表している。

さて，原問題(1)の最大期待値は $v^0(x; \lambda)$ に $(x_0; 1)$ を代入した値 $v^0(x_0; 1)$ で与えられる。また，最大点（点とはいえ，ここでは確率過程）は拡大履歴空間上の最適行動(7)を元の履歴空間 $H=X\times U\times X\times U\times X$ 上に射影した行動

$$x_0\rightarrow u_0^*\rightarrow X_1^*\rightarrow U_1^*\rightarrow X_2^* \tag{10}$$

で与えられる。これが本来の最適行動である。すなわち，各段評価の最悪値をシステム全体としての評価値としたとき，(10)がその期待値を最大にする行動である。この最適行動は推移確率 $p_0^*(x_1|x_0)$, $p_1^*(x_2|x_0, x_1)$ をもつ X 上の確率過程になっている。ただし

$$p_0^*(x_1|x_0)=p(x_1|x_0, u_0^*), \quad p_1^*(x_2|x_0, x_1)=p(x_2|x_1, u_1^*) \tag{11}$$

ここに

$$u_0^*=\pi_0^*(x_0; 1), \quad \lambda_1=r_0(u_0^*), \quad u_1^*=\pi_1^*(x_1; \lambda_1) \tag{12}$$

だから，u_0^*は x_0 に依存し，u_1^*は x_0, x_1 の関数になっている。

さて，ここで注意すべきは，本来の制御マルコフ連鎖上で最大期待値に到達する（最適な）確率過程 $\{x_0, X_1^*, X_2^*\}$ は必ずしも X 上のマルコフ連鎖になっていないということである。すなわち，原問題(1)の最適政策はマルコフ政策の中には存在しないのである。事実，最適過程 $\{x_0, X_1^*, X_2^*\}$ を生じせしめる最適政策 $\sigma^*=\{\sigma_0^*, \sigma_1^*\}$ の第2決定関数は初期状態 x_0 にも依存していて，X 上のマルコフ政策になっていない。原問題(1)の最適政策を構成する決定関数の表現は拡大過程(6)の最適政策 $\pi^*=\{\pi_0^*, \pi_1^*\}$ を用いて

$$\sigma_0^*(x_0)=\pi_0^*(x_0; 1), \quad \sigma_1^*(x_0, x_1)=\pi_1^*(x_1; \lambda_1) \tag{13}$$

で与えられる。ここでも λ_1 は r_0, u_0^* を通して x_0 の**関数**になっていることに注意すべきである。

5．埋め込まない場合

他方，パラメータ λ を導入するような不変埋没原理を用いないで，加法型評価系の問題と同様に部分問題群

$$\begin{aligned} v^2(x_2)&=r_G(x_2) \\ v^1(x_1)&=\underset{\pi_1}{\mathrm{Max}}\sum_{x_2\in X}[r_1(u_1)\wedge r_G(x_2)]p(x_2|x_1, u_1) \\ v^0(x_0)&=\underset{\pi_0,\pi_1}{\mathrm{Max}}\sum_{(x_1,x_2)\in X\times X}\{[r_0(u_0)\wedge r_1(u_1)\wedge r_G(x_2)] \\ &\qquad\times p(x_1|x_0, u_0)p(x_2|x_1, u_1)\} \end{aligned} \tag{14}$$

を定義することもできる。$v^1(x_1)$ では $u_1=\pi_1(x_1)$ であり，第2決定関数 π_1 を動かして最大化する。また，$v^0(x_0)$ では $u_0=\pi_1(x_0)$，$u_1=\pi_1(x_1)$ である。したがって，$v^0(x_0)$ での最大化は政策 $\pi=\{\pi_0, \pi_1\}$ 全体で取る。

しかしこのとき，一般にこの1変数最大値関数列 $\{v^0, v^1, v^2\}$ 間には再帰式

$$
\begin{aligned}
v^2(x) &= r_G(x) \\
v^1(x) &= \operatorname*{Max}_{u\in U}\Big[r_1(u)\wedge \sum_{y\in X} v^2(y)p(y|x, u)\Big] \\
v^0(x) &= \operatorname*{Max}_{u\in U}\Big[r_0(u)\wedge \sum_{y\in X} v^1(y)p(y|x, u)\Big]
\end{aligned}
\tag{15}
$$

は成立しない[2)]ということに注意しなければならない([1], [4], [5], [11], [14], [15], [18] を参照)。もちろん，分離等式

$$
v^n(x;\lambda)=v^n(x)\wedge\lambda \tag{16}
$$

も一般には成り立たない。

6．最大過程上の終端評価

さて，X 上の最小型評価である原問題(1)を埋め込んだ問題群(5)は拡大状態空間 $Y=X\times[0, 1]$ 上で終端型評価をもつ次の決定過程に同値変換されることを見よう：

$$
\begin{aligned}
&\text{Maximize} && \tilde{E}[R(y_2)] \\
&\text{subject to} && \text{(i)}\ \ y_{n+1}\sim q_n(\cdot|y_n, u_n) \qquad n=0, 1 \\
& && \text{(ii)}\ \ u_0\in U,\ u_1\in U
\end{aligned}
\tag{17}
$$

ただし，マルコフ推移確率 $q_0(y_1|y_0, u_0)$，$q_1(y_2|y_1, u_1)$ および終端評価関数 R はそれぞれ

$$
\begin{aligned}
q_0((y,\mu)|(x,\lambda), u) &= \begin{cases} p(y|x, u) & \mu=\lambda\wedge r_0(u)\ \text{のとき} \\ 0 & \text{その他} \end{cases} \\
q_1((y,\mu)|(x,\lambda), u) &= \begin{cases} p(y|x, u) & \mu=\lambda\wedge r_1(u)\ \text{のとき} \\ 0 & \text{その他} \end{cases}
\end{aligned}
\tag{18}
$$

$$
R(x,\lambda)=\lambda\wedge r_G(x) \tag{19}
$$

で定まる。このとき，期待値作用素 $\tilde{E}$ は，初期状態 y_0，マルコフ推移確率 q_0, q_1 および政策 $\pi=\{\pi_0, \pi_1\}$ から履歴の直積空間 $\tilde{H}$ 上に唯一定まる確率測度による。

この拡大過程について部分問題群を定義して再帰式を導こう。各 $y_n=(x_n, \lambda)\in Y$, $n=0, 1, 2$ に対して部分問題群

$$
\begin{aligned}
w^2(y_2) &= R(y_2) \\
w^1(y_1) &= \operatorname*{Max}_{\pi_1}\sum_{y_2\in Y} R(y_2)q_1(y_2|y_1, u)
\end{aligned}
\tag{20}
$$

2) 付録3で反例を与える。

$$w^0(y_0)=\underset{\pi_0,\pi_1}{\mathrm{Max}}\sum_{(y_1,y_2)\in Y\times Y}R(y_2)q_0(y_1|y_0,\ u_0)q_1(y_2|y_1,\ u_1) \tag{21}$$

を定義する。ただし，上式(20), (21)における最大化は拡大状態空間 $Y=X\times[0,\ 1]$ 上での決定関数 $\pi_0,\ \pi_1：Y\to U$ 全体にわたるものとする。言い換えれば，$w^1(y_1)$ における最大化は第2決定関数 π_1 を動かし，$w^0(y_0)$ では政策 $\pi=\{\pi_0,\ \pi_1\}$ 全体で取る。このとき，再帰式

定理 2 （[9]）

$$\begin{aligned} w^2(y)&=R(y)\\ w^1(y)&=\underset{u\in U}{\mathrm{Max}}\sum_{z\in Y}w^2(z)q_1(z|y,\ u)\\ w^0(y)&=\underset{u\in U}{\mathrm{Max}}\sum_{z\in Y}w^1(z)q_0(z|y,\ u)\end{aligned} \tag{22}$$

が成立する。ここで $y=(x,\lambda)$，推移法則(18)および終端関数(19)を考慮して，(x,λ) で表すと，再帰式(22)は

$$\begin{aligned} w^2(x,\lambda)&=\lambda\wedge r_G(x)\\ w^1(x,\lambda)&=\underset{u\in U}{\mathrm{Max}}\sum_{y\in X}w^2(y,\lambda\wedge r_1(x))p(y|x,\ u)\\ w^0(x,\lambda)&=\underset{u\in U}{\mathrm{Max}}\sum_{y\in X}w^1(y,\lambda\wedge r_0(x))p(y|x,\ u)\end{aligned} \tag{23}$$

になる。二つの再帰式(6)と(23)を比較すれば，

$$v^n(x;\lambda)=w^n(x,\lambda)\qquad n=0,\ 1,\ 2 \tag{24}$$

が分かる。すなわち，最適値関数としてみれば，埋め込まれた問題の部分問題群(4)と拡大過程の部分問題群(22)は同値である。

7．定式化と最適性の再検討

これまで，X 上の最小型評価をもつ原問題(1)を拡大状態空間 $Y=X\times[0,1]$ 上で終端型評価をもつ決定過程(17)に同値変換して，拡大過程(17)の最適政策をマルコフ政策の枠組みの中で求めて，それを元の空間に射影して(1)の最適解を求めてきた。

この節では，原問題(1)の定式化そのものを含めて，(1)の最適性を再考してみよう。さて，原問題(1)における期待値作用素 E は初期状態 x_0 のみならず政策 π にも依存するので，$E^{\pi}_{x_0}$ であった。その意味で厳密には原問題(1)は

$$\begin{array}{ll}\text{Maximize} & E^{\pi}_{x_0}[r_0(u_0)\wedge r_1(u_1)\wedge r_G(x_2)]\\ \text{subject to} & \text{(i)}\ \ x_{n+1}\sim p(\cdot|x_n,\ u_n)\qquad n=0,\ 1\\ & \text{(ii)}\ \ u_0\in U,\quad u_1\in U\end{array} \tag{25}$$

で表す方がより適切である。この最大化は X 上のマルコフ政策 $\pi=\{\pi_0,\ \pi_1\}$ 全体についてとるものであった。すなわち，ここでは決定関数 $\pi_0,\ \pi_1$ はそれぞれ X 上の関数

$$\pi_0：X\to U,\quad \pi_1：X\to U \tag{26}$$

であった。もっと厳密には，**原問題(1)は2つの式(25), (26)で表されている**というべきである。しか

し，(13)から分かるように，拡大過程(4)から導かれた最適政策 $\sigma^*=\{\sigma_0^*, \sigma_1^*\}$ の決定関数列は

$$\sigma_0^*: X \to U, \quad \sigma_1^*: X \times X \to U \tag{27}$$

であった。この第2決定関数は初期状態 x_0 にも依存しているので，最適政策 $\sigma^*=\{\sigma_0^*, \sigma_1^*\}$ は X 上のマルコフ政策になっていない。すなわち，原問題(1)の**最適政策は一般にマルコフ政策にならない。**

これは最初の定式化に**矛盾**があることを示している。この矛盾を回避するために，原問題を修正してあらためて期待値最適化問題

$$\begin{array}{lll} \text{Maximize} & E_{x_0}^{\sigma}[r_0(u_0) \wedge r_1(u_1) \wedge r_G(x_2)] & \\ \text{subject to} & \text{(i)} \quad x_{n+1} \sim p(\cdot|x_n, u_n) \qquad n=0, 1 & \\ & \text{(ii)} \quad u_0 \in U, \quad u_1 \in U & \end{array} \tag{28}$$

の最大化をマルコフ政策を含む**一般の**政策 $\sigma=\{\sigma_0, \sigma_1\}$ 全体についてとるものを考える。すなわち，ここでは決定関数 σ_0 は X 上の，そして σ_1 は $X \times X$ 上のそれぞれ関数である：

$$\sigma_0: X \to U, \quad \sigma_1: X \times X \to U. \tag{29}$$

このように**修正された問題**(28)，(29)に対しては部分問題群

$$V^2(x_2) = r_G(x_2)$$

$$V^1(x_1) = \operatorname*{Max}_{\sigma_1} \sum_{x_2 \in X} [r_1(u_1) \wedge r_G(x_2)] p(x_2|x_1, u_1) \tag{30}$$

$$\begin{aligned} V^0(x_0) = \operatorname*{Max}_{\sigma_0, \sigma_1} \sum_{x_1, x_2 \in X \times X} \{ & [r_0(u_0) \wedge r_1(u_1) \wedge r_G(x_2)] \\ & \times p(x_1|x_0, u_0) p(x_2|x_1, u_1)\} \end{aligned} \tag{31}$$

を定義することができる。ただし，式(30)での最大化は第2決定関数 $\sigma_1: X \to U$ 全体で取り，式(31)では(マルコフとはかぎらない)(29)の政策全体で取るものである。すなわち，$V^1(x_1)$ では $u_1=\sigma_1(x_1)$ であり，第2決定関数 σ_1 を動かして最大化する。また，$V^0(x_0)$ では $u_0=\sigma_1(x_0)$，$u_1=\sigma_1(x_0, x_1)$ である。したがって，$V^0(x_0)$ での最大化は政策 $\sigma=\{\sigma_0, \sigma_1\}$ 全体で取る。

さて，パラメータを含まない部分問題群(14)と比較すれば，明らかに

$$V^2(x_2) = v^2(x_2), \quad V^1(x_1) = v^1(x_1), \quad V^0(x_0) \geq v^0(x_0) \tag{32}$$

である。しかしここでも，この1変数最大値関数列 $\{V^0, V^1, V^2\}$ 間には再帰式

$$\begin{aligned} V^2(x) &= r_G(x) \\ V^1(x) &= \operatorname*{Max}_{u \in U} [r_1(u) \wedge \sum_{y \in X} V^2(y) p(y|x, u)] \\ V^0(x) &= \operatorname*{Max}_{u \in U} [r_0(u) \wedge \sum_{y \in X} V^1(y) p(y|x, u)] \end{aligned} \tag{33}$$

は依然として一般には成立しない[3]ということに注意しなければならない。

いずれにしても，埋め込まないままで原問題を解こうとすると，いわゆる再帰式を導く段階で（最小型関数の期待値最大化ゆえ）無理が生じる。この点，評価値の累積度を表す1次元パラメータ λ を導入して，状態空間を拡大した決定過程を考えることが極めて重要である。この拡大過程のマルコフ

3)　付録3で反例を与える。

政策に関する再帰式を導いて，これを解いて得た最適解を原問題に射影して，求める最適解を得ることができるからである。このように，不変埋没原理に基づいて最小型関数の期待値最大化問題を解くことは避けられない方法である。この方法を**確率的終端状態接近方法**（*stochastic terminal state approach*）と呼ぼう。いわゆる終端状態接近方法（terminal state approach）については Sniedovich［17, p.211］が final state scheme の名の下で論じているが，その対象は**確定的最適化問題に限定**されていた。この意味においてこれまで用いられてきた方法は**確定的**終端状態接近方法（*deterministic* terminal state approach）であると言える。しかし，確率的問題では最適政策はマルコフ政策でないなどの特異な状況が生じ，確定的問題のアプローチはそのまま使えなくなっている。したがって，本論文では新しく**確率的**終端状態接近方法を導入して，これまで取り扱われていなかった最小型関数の期待値最大化問題を解決したことになる。

8．おわりに

本論文では，不確実性の下での逐次意志決定過程を最小型評価関数の期待値最大化問題としてとらえることによって，**最適政策は必ずしもマルコフ政策でない**ことを新しい数学的最適化手法を創り出して証明した。この事実は，現実の不透明な状況に対する最適な危機管理政策は「今だけをみて行動する」ことではないことを示している。その理由の一つは評価型の最小性にあり，また不確実性からくることでもある。例えば，初期の大地震による経験・教訓は後世まで引き継がれていくべきであることを示している。また，いわゆる危機対策としては，初期からの対策が重要であることを示している。すなわち，初動捜査を含めた継続捜査の重要性を説いている。

結論的には，多期間にわたる不確実な状況の下で危険，危機，後悔，ショックというものを総体として和らげる**最適な方法は，**単に現在だけを見て行動するのではなく，**過去から現時点までの累積評価値を絶えず見据えて行動することである**ことを示している。

A．付録1：加法型過程

これまでの本論文の議論において，最小演算記号 $\wedge$ をすべて加法演算記号 $+$ に置き換えることによって，**加法型過程**—加法型関数の期待値最大化問題—が得られる。加法型過程が危機管理問題等を表現しているとは到底考えられないが，この加法型過程についてはこれまでと同様に最適値

$$v^n(x_n;\lambda),\quad v^n(x_n),\quad w^n(y_n),\quad V^n(x_n)$$

が定義される（[2]，[6]，[16]，[19]）。しかも数学的にはこれまでの議論に出てきた矛盾など起こることなく，これらの量の間には望ましい等式関係が成立する。例えば，

定理3

$$v^2(x;\lambda)=\lambda+r_G(x)$$

$$v^1(x;\lambda)=\underset{u\in U}{\mathrm{Max}}\sum_{y\in X}v^2(y;\lambda+r_1(u))p(y|x,u) \tag{34}$$
$$v^0(x;\lambda)=\underset{u\in U}{\mathrm{Max}}\sum_{y\in X}v^1(y;\lambda+r_0(u))p(y|x,u)$$

$$\begin{aligned} v^2(x)&=r_G(x)\\ v^1(x)&=\underset{u\in U}{\mathrm{Max}}[r_1(u)+\sum_{y\in X}v^2(y)p(y|x,u)]\\ v^0(x)&=\underset{u\in U}{\mathrm{Max}}[r_0(u)+\sum_{y\in X}v^1(y)p(y|x,u)]\end{aligned} \tag{35}$$

が成り立ち，当然分離等式

$$v^n(x;\lambda)=v^n(x)+\lambda \tag{36}$$

は成立する。また，拡大過程の再帰式

定理 4

$$\begin{aligned} w^2(y)&=R(y)\\ w^1(y)&=\underset{u\in U}{\mathrm{Max}}\sum_{z\in Y}w^2(z)q_1(z|y,u)\\ w^0(y)&=\underset{u\in U}{\mathrm{Max}}\sum_{z\in Y}w^1(z)q_0(z|y,u)\end{aligned} \tag{37}$$

も成り立って，(x,λ) で表すと，

$$\begin{aligned} w^2(x,\lambda)&=\lambda+r_G(x)\\ w^1(x,\lambda)&=\underset{u\in U}{\mathrm{Max}}\sum_{y\in X}w^2(y,\lambda+r_1(x))p(y|x,u)\\ w^0(x,\lambda)&=\underset{u\in U}{\mathrm{Max}}\sum_{y\in X}w^1(y,\lambda+r_0(x))p(y|x,u)\end{aligned} \tag{38}$$

になる。結局，ここでも

$$v^n(x;\lambda)=w^n(x,\lambda) \qquad n=0,1,2 \tag{39}$$

になる。さらに，等式

$$V^2(x_2)=v^2(x_2),\quad V^1(x_1)=v^1(x_1),\quad V^0(x_0)=v^0(x_0) \tag{40}$$

が成立する。したがって，再帰式

定理 5

$$\begin{aligned} V^2(x)&=r_G(x)\\ V^1(x)&=\underset{u\in U}{\mathrm{Max}}[r_1(u)+\sum_{y\in X}V^2(y)p(y|x,u)]\\ V^0(x)&=\underset{u\in U}{\mathrm{Max}}[r_0(u)+\sum_{y\in X}V^1(y)p(y|x,u)]\end{aligned} \tag{41}$$

も成立する。さらに，§4 最適解の構成をこの加法型過程に対して同じ流れで，しかし注意深く議論していくと，最適政策 $\sigma^*=\{\sigma_0^*(x_0),\sigma_1^*(x_0,x_1)\}$ の第 2 決定関数 $\sigma_1^*(x_0,x_1)$ はその第 1 (初期) 状態 x_0 に依存しないことが分かる。すなわち，**加法型過程では最適政策はマルコフ政策になる**。このことが**最小型過程**―最小型関数の期待値最大化問題―との決定的な違いである。

最後に，本論文§2 問題と定式化において，最小型過程をいわゆるマルコフ政策の枠組みの中で定式化したことに言及しておこう。既に30年以上におよぶマルコフ決定過程研究の結果から，最適政策を求めるには（加法型過程の場合だが!!）**マルコフ政策に限定してよい**(Markov policy is enough) と

いう事実ある([2],[6],[16],[19])。実は，§2の定式化はこの事実を前提としたのである。しかし，マルコフ政策の中での定式化は矛盾をきたし，最小型過程の場合は一般政策の枠組みの中で定式化すると，矛盾は解決することが分かった。このように，最小型関数の期待値最大化問題をマルコフ政策の中で定式化することによって，既存の加法型関数の期待値最適化との相違が鮮やかに浮き彫り出来たことになる。そして，加法型過程の研究成果を携えて，いま最小型関数を含む結合型関数の期待値最適化の研究は緒についたと言えよう([9],[10])。

B. 付録2：証明

さて，パラメータλを含む部分問題群

$$v^2(x_2;\lambda)=\lambda\wedge r_G(x_2)$$

$$v^1(x_1;\lambda)=\underset{\pi_1}{\mathrm{Max}}\sum_{x_2\in X}[\lambda\wedge r_1(u_1)\wedge r_G(x_2)]p(x_2|x_1,u_1) \tag{42}$$

$$v^0(x_0;\lambda)=\underset{\pi_0,\pi_1}{\mathrm{Max}}\sum_{(x_1,x_2)\in X\times X}\{[\lambda\wedge r_0(u_0)\wedge r_1(u_1)\wedge r_G(x_2)] \times p(x_1|x_0,u_0)p(x_2|x_1,u_1)\} \tag{43}$$

の式(42)においては$u_1=\pi_1(x_1;\lambda)$であり，式(43)では$u_0=\pi_0(x_0;\lambda)$，$u_1=\pi_1(x_1;\lambda_1)$であることに注意する。ただし，$\lambda_1=\lambda\wedge r_0(u_0)$.

このとき，2変数最大値関数列$\{v^0,v^1,v^2\}$が再帰式

$$v^2(x;\lambda)=\lambda\wedge r_G(x) \tag{44}$$

$$v^1(x;\lambda)=\underset{u\in U}{\mathrm{Max}}\sum_{y\in X}v^2(y;\lambda\wedge r_1(u))p(y|x,u) \tag{45}$$

$$v^0(x;\lambda)=\underset{u\in U}{\mathrm{Max}}\sum_{y\in X}v^1(y;\lambda\wedge r_0(u))p(y|x,u) \tag{46}$$

を満たすことを示す。

証明　最初の2つの式(44),(45)は明らかである。第3の式(46)すなわち

$$v^0(x_0;\lambda_0)=\underset{u_0\in U}{\mathrm{Max}}\sum_{x_1\in X}v^0(x_1;\lambda_0\wedge r_0(x_0,u_0))p(x_1|x_0,u_0)\quad x_0\in X,\quad \lambda_0\in[0,1] \tag{47}$$

の成立を示す。

まず，1段過程の最適(マルコフ)政策π_1^*を選ぶと，

$$v^1(x_1;\lambda_1)=\sum_{x_2\in X}[\lambda_1\wedge r_1(u_1)\wedge r_G(x_2)]p(x_2|x_1,u_1)\quad \forall(x_1,\lambda_1)\in X\times[0,1] \tag{48}$$

$$(u_1=\pi_1^*(x_1;\lambda_1))$$

になる。2段過程の定義(5)より，各$(x_0,\lambda_0)\in X\times[0,1]$において最適（マルコフ）政策$\tilde{\pi}=\{\tilde{\pi}_0,\tilde{\pi}_1\}$を選ぶと，

$$v^0(x_0;\lambda_0)=\sum_{(x_1,x_2)\in X\times X}\{[\lambda_0\wedge r_0(u_0)\wedge r_1(u_1)\wedge r_G(x_2)] \times p(x_1|x_0,u_0)p(x_2|x_1,u_1)\} \tag{49}$$

になる。ただし

$$u_0 = \tilde{\pi}_0(x_0; \lambda_0), \quad \lambda_1 = \lambda_0 \wedge r_0(u_0), \quad u_1 = \tilde{\pi}_1(x_1; \lambda_1). \tag{50}$$

以下では普遍式

$$\sum\sum_{(x_1, x_2) \in X \times X} f(x_1, x_2) = \sum_{x_1 \in X} \sum_{x_2 \in X} f(x_1, x_2) \tag{51}$$

も用いる。また，不等式

$$\begin{aligned} &\sum_{x_2 \in X} [\lambda_1 \wedge r_1(u_1) \wedge r_G(x_2)] p(x_2|x_1, u_1) \\ &\leq v^1(x_1; \lambda_1) \quad \forall (x_1, \lambda_1) \in X \times [0, 1] \end{aligned} \tag{52}$$

が成り立つ。したがって，式(49)～(52)より，次が成り立つ：

$$\begin{aligned} v^0(x_0; \lambda_0) &\leq \sum_{x_1 \in X} \sum_{x_2 \in X} \{[\lambda_0 \wedge r_0(u_0) \wedge r_1(u_1) \wedge r_G(x_2))] \\ &\qquad\qquad \times p(x_1|x_0, u_0) p(x_2|x_1, u_1)\} \\ &= \sum_{x_1 \in X} \sum_{x_2 \in X} \{[\lambda_1 \wedge r_1(u_1) \wedge r_G(x_2)] p(x_2|x_1, u_1)\} p(x_1|x_0, u_0)\} \\ &\qquad\qquad (\lambda_1 = \lambda_0 \wedge r_0(u_0)) \\ &\leq \sum_{x_1 \in X} v^1(x_1; \lambda_1) p(x_1|x_0, u_0) \qquad (\lambda_1 = \lambda_0 \wedge r_0(u_0)) \\ &= \sum_{x_1 \in X} v^1(x_1; \lambda_0 \wedge r_0(u_0)) p(x_1|x_0, u_0). \end{aligned} \tag{53}$$

よって，

$$v^0(x_0; \lambda_0) \leq \sum_{x_1 \in X} v^1(x_1; \lambda_0 \wedge r_0(u_0)) p(x_1|x_0, u_0) \qquad \forall (x_0, \lambda_0) \in X \times [0, 1] \tag{54}$$

である。ここで，すべての $u \in U$ について最大化演算を施すと，不等式

$$v^0(x_0; \lambda_0) \leq \operatorname*{Max}_{u_0 \in U} \sum_{x_1 \in X} v^1(x_1; \lambda_0 \wedge r_0(u_0)) p(x_1|x_0, u_0) \quad \forall (x_0, \lambda_0) \in X \times [0, 1] \tag{55}$$

が得られる。

他方，任意の $(x_0, \lambda_0) \in X \times [0, 1]$ に対して $u^* = u^*(x_0, \lambda_0) \in U$ が式(55)の右辺の最大値に到達するとすれば，マルコフ決定関数

$$\pi_0^* : X \times [0, 1] \to U \quad \pi_0^*(x_0; \lambda_0) = u^*$$

が定義できて，

$$\begin{aligned} &\operatorname*{Max}_{u_0 \in U} \sum_{x_1 \in X} v^1(x_1; \lambda_0 \wedge r_0(u_0)) p(x_1|x_0, u_0) \\ &= \sum_{x_1 \in X} v^1(x_1; \lambda_0 \wedge r_0(u_0)) p(x_1|x_0, u_0) \quad (u_0 = \pi_0^*(x_0; \lambda_0)) \end{aligned} \tag{56}$$

である。したがって，(48)より

$$\begin{aligned} &\sum_{x_1 \in X} v^1(x_1; \lambda_0 \wedge r_0(u_0)) p(x_1|x_0, u_0) \quad (u_0 = \pi_0^*(x_0; \lambda_0)) \\ &= \sum_{x_1 \in X} \{ \sum_{x_2 \in X} [\lambda_1 \wedge r_1(u_1) \wedge r_G(x_2)] p(x_2|x_1, u_1)\} p(x_1|x_0, u_0) \\ &\qquad\qquad (\lambda_1 = \lambda_0 \wedge r_0(u_0)) \\ &= \sum\sum_{(x_1, x_2) \in X \times X} \{[\lambda_0 \wedge r_0(u_0) \wedge r_1(u_1) \wedge r_G(x_2)] p(x_1|x_0, u_0) p(x_2|x_1, u_1)\} \end{aligned} \tag{57}$$

になる。更に，式(56), (57)より，

$$\operatorname*{Max}_{u_0 \in U} \sum_{x_1 \in X} v^1(x_1; \lambda_0 \wedge r_0(u_0)) p(x_1|x_0, u_0)$$

$$
\begin{aligned}
&= \sum\sum_{(x_1,x_2)\in X\times X} \{[\lambda_0 \wedge r_0(u_0) \wedge r_1(u_1) \wedge r_G(x_2)]p(x_1|x_0, u_0)p(x_2|x_1, u_1)\} \\
&\qquad\qquad (u_0=\pi_0^*(x_0; \lambda_0), \quad \lambda_2=\lambda_0\wedge r_0(u_0), \quad u_1=\pi_1^*(x_1; \lambda_1)) \\
&\le \operatorname*{Max}_{\pi_0,\pi_1} \sum\sum_{(x_1,x_2)\in X\times X} \{[\lambda_0 \wedge r_0(u_0) \wedge r_1(u_1) \wedge r_G(x_2)]p(x_1|x_0, u_0)p(x_2|x_1, u_1)\} \\
&= v^0(x_0; \lambda_0)
\end{aligned} \tag{58}
$$

になる。すなわち，式(55)と(58)より，等式(47)の成立が示された。(証明終り)

なお，この等式の証明には拡大過程に対する終端型評価系において確率的な意味での単調性

$$
b_i \ge c_i(i=1, \cdots, n) \Longrightarrow \sum_{i=1}^{n}(a\circ b_i)p_i \ge \sum_{i=1}^{n}(a\circ c_i)p_i \tag{59}
$$

$$
(a\circ b=b, \quad \sum_{i=1}^{n} p_i=1, \quad p_i\ge 0)
$$

を用いたことになるが，確率的最適化である原過程の最小型評価系に対しては確定的な意味での単調性

$$
\operatorname*{Max}_{x}[c\wedge f(x)]=c\wedge \operatorname*{Max}_{x} f(x) \tag{60}
$$

は用いていないことに注意しよう。

C　付録3：脚注

ここでは脚注2，脚注3に同時に答える反例を示す。

1変数最大値関数列 $\{v^0, v^1, v^2\}$ 間には再帰式

$$
\begin{aligned}
v^2(x) &= r_G(x) \\
v^1(x) &= \operatorname*{Max}_{u\in U}[r_1(u) \wedge \sum_{y\in X} v^2(y)p(y|x, u)] \\
v^0(x) &= \operatorname*{Max}_{u\in U}[r_0(u) \wedge \sum_{y\in X} v^1(y)p(y|x, u)]
\end{aligned} \tag{61}
$$

は一般には成立しない(脚注2)。

1変数最大値関数列 $\{V^0, V^1, V^2\}$ 間には再帰式

$$
\begin{aligned}
V^2(x) &= r_G(x) \\
V^1(x) &= \operatorname*{Max}_{u\in U}[r_1(u) \wedge \sum_{y\in X} V^2(y)p(y|x, u) \\
V^0(x) &= \operatorname*{Max}_{u\in U}[r_0(u) \wedge \sum_{y\in X} V^1(y)p(y|x, u)
\end{aligned} \tag{62}
$$

は一般には成立しない(脚注3)。

ここでは次のような Bellman and Zadeh［4, pp. B154］の数値例からなる2段過程を考えよう。

$$
r_G(s_1)=0.3, \quad r_G(s_2)=1.0, \quad r_G(s_3)=0.8 \tag{63}
$$

$$
r_1(a_1)=1.0, \qquad r_1(a_2)=0.6 \tag{64}
$$

$$
r_0(a_1)=0.7, \qquad r_0(a_2)=1.0 \tag{65}
$$

$u_t = a_1$

$x_t \diagdown x_{t+1}$	s_1	s_2	s_3
s_1	0.8	0.1	0.1
s_2	0.0	0.1	0.9
s_3	0.8	0.1	0.1

$u_t = a_2$

$x_t \diagdown x_{t+1}$	s_1	s_2	s_3
s_1	0.1	0.9	0.0
s_2	0.8	0.1	0.1
s_3	0.1	0.0	0.9

まず，

$$V^2(s_1)=0.3,\quad V^2(s_2)=1.0,\quad V^2(s_3)=0.8 \tag{66}$$

は明かである。次に，最大期待値 $\{V^1(s_1),\ V^1(s_2),\ V^1(s_3)\}$ は以下の全数直接列挙法（図１）によって

$$V^1(s_1)=0.57,\quad V^1(s_2)=0.82,\quad V^1(s_3)=0.57 \tag{67}$$

であることがわかる。同時に，最大値に到達する値を探すと，最適（当然マルコフ）政策 $\{\pi_1^*(x_1)\}$

$$\pi_1^*(s_1)=a_2,\quad \pi_1^*(s_2)=a_1,\quad \pi_1^*(s_3)=a_2 \tag{68}$$

が得られる。これは図１の実線に対応している。

図１では，次の簡略記号を用いている：

履歴＝$x_1 \quad r_1(u_1)/u_1 \quad p(x_2|x_1, u_1) \quad x_2$

終端＝終端評価値＝$r_G(x_2)$

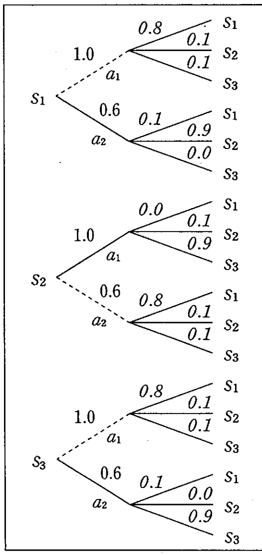
経路＝経路確率＝$p(x_2|x_1, u_1)$

最小＝最小評価値＝$r_1(u_1) \wedge r_G(x_2)$

積＝経路×最小

さらに，イタリック体は確率を，**ボールド体**は上下の期待値の最大（大きい方）の数値を表す。

図１　１段過程に対する状態 s_1, s_2, s_3 からの全行動と最大分枝の選択



終端	経路	最小	積	期待値
0.3	*0.8*	0.3	0.24	
1.0	*0.1*	1.0	0.1	0.42
0.8	*0.1*	0.8	0.08	
0.3	*0.1*	0.3	0.03	
1.0	*0.9*	0.6	0.54	**0.57**
0.8	*0.0*	0.6	0.0	
0.3	*0.0*	0.3	0.0	
1.0	*0.1*	1.0	0.1	**0.82**
0.8	*0.9*	0.8	0.72	
0.3	*0.8*	0.3	0.24	
1.0	*0.1*	0.6	0.06	0.36
0.8	*0.1*	0.6	0.06	
0.3	*0.8*	0.3	0.24	
1.0	*0.1*	1.0	0.1	0.42
0.8	*0.1*	0.8	0.08	
0.3	*0.1*	0.3	0.03	
1.0	*0.0*	0.6	0.0	**0.57**
0.8	*0.9*	0.6	0.54	

最後に，２段過程の最大期待値 $\{V^0(s_1),\ V^0(s_2),\ V^0(s_3)\}$ は以下の図２，３，４による全行動列挙法によって

$$V^0(s_1)=0.795,\quad V^0(s_2)=0.595,\quad V^0(s_3)=0.583 \tag{69}$$

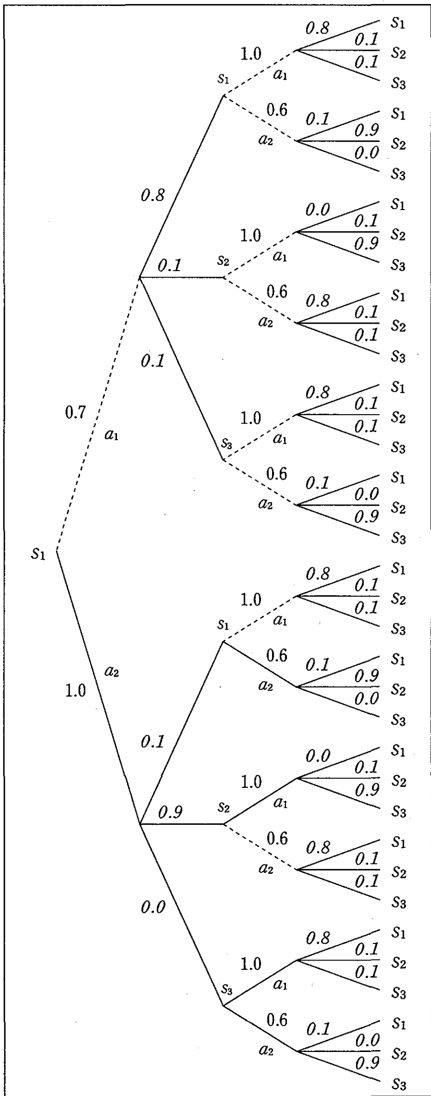
になる。このとき，最大期待値の上下選択によって最適政策

$$\pi^*=\{\pi_0^*(x_0),\ \pi_1^*(x_0,x_1)\}$$

は

$$\pi_0^*(s_1)=a_2,\quad \pi_0^*(s_2)=a_2,\quad \pi_0^*(s_3)=a_1 \tag{70}$$

図２　２段過程に対する状態 s_1 からの全行動と最大分枝の選択



履歴	終端	経路	最小	積	部分期	全期待
	0.3	0.64	0.3	0.192		
	1.0	0.08	0.7	0.056	0.304	
	0.8	0.08	0.7	0.056		
	0.3	0.08	0.3	0.024		
	1.0	0.72	0.6	0.432	**0.456**	
	0.8	0.0	0.6	0.0		
	0.3	0.0	0.3	0.0		
	1.0	0.01	0.7	0.007	**0.070**	
	0.8	0.09	0.7	0.063		0.583
	0.3	0.08	0.3	0.024		
	1.0	0.01	0.6	0.006	0.036	
	0.8	0.01	0.6	0.006		
	0.3	0.08	0.3	0.024		
	1.0	0.01	0.7	0.007	0.038	
	0.8	0.01	0.7	0.007		
	0.3	0.01	0.3	0.003		
	1.0	0.0	0.6	0.0	**0.057**	
	0.8	0.09	0.6	0.054		
	0.3	0.08	0.3	0.024		
	1.0	0.01	1.0	0.01	0.042	
	0.8	0.01	0.8	0.008		
	0.3	0.01	0.3	0.003		
	1.0	0.09	0.6	0.054	**0.057**	
	0.8	0.0	0.6	0.0		
	0.3	0.0	0.3	0.0		
	1.0	0.09	1.0	0.09	**0.738**	
	0.8	0.81	0.8	0.648		**0.795**
	0.3	0.72	0.3	0.216		
	1.0	0.09	0.6	0.054	0.324	
	0.8	0.09	0.6	0.054		
	0.3	0.0	0.3	0.0		
	1.0	0.0	1.0	0.0	**0.0**	
	0.8	0.0	0.8	0.0		
	0.3	0.0	0.3	0.0		
	1.0	0.0	0.6	0.0	**0.0**	
	0.8	0.0	0.6	0.0		

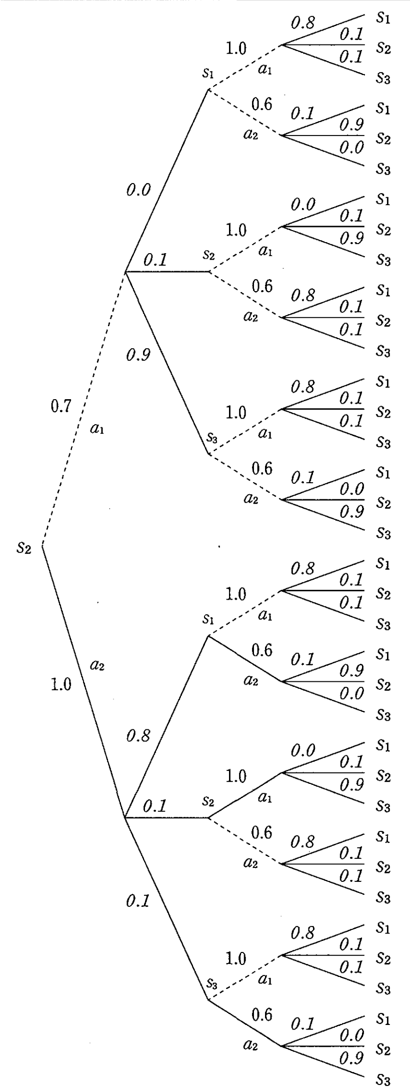
$$\pi_0^*(s_1, s_1)=a_2, \quad \pi_0^*(s_1, s_2)=a_1, \quad \pi_0^*(s_1, s_3)=a_2$$
$$\pi_0^*(s_2, s_1)=a_2, \quad \pi_0^*(s_2, s_2)=a_1, \quad \pi_0^*(s_2, s_3)=a_2 \tag{71}$$
$$\pi_0^*(s_3, s_1)=a_2, \quad \pi_0^*(s_3, s_2)=a_1, \quad \pi_0^*(s_3, s_3)=a_2$$

で表されることがわかる。ここで注意すべきは第 2 決定関数 $\pi_1^*(x_0, x_1)$ は初期状態 x_0 に依存していないということである：

$$\pi_1^*(x_0, x_1)=\pi_1^*(x_1).$$

したがって，最適政策 π^* はマルコフ政策であることがわかる。故に，この事実より等式 $v^0(x_0)=$

図 3　2 段過程に対する状態 s_2 からの全行動と最大分枝の選択



履　歴	終端	経路	最小	積	部分期	全期待
s_1	0.3	*0.0*	0.3	0.0		
s_2	1.0	*0.0*	0.7	0.0	**0.0**	
s_3	0.8	*0.0*	0.7	0.0		
s_1	0.3	*0.0*	0.3	0.0		
s_2	1.0	*0.0*	0.6	0.0	**0.0**	
s_3	0.8	*0.0*	0.6	0.0		
s_1	0.3	*0.0*	0.3	0.0		
s_2	1.0	*0.01*	0.7	0.007	**0.070**	
s_3	0.8	*0.09*	0.7	0.063		
s_1	0.3	*0.08*	0.3	0.024		0.583
s_2	1.0	*0.01*	0.6	0.006	0.036	
s_3	0.8	*0.01*	0.6	0.006		
s_1	0.3	*0.72*	0.3	0.216		
s_2	1.0	*0.09*	0.7	0.063	0.342	
s_3	0.8	*0.09*	0.7	0.063		
s_1	0.3	*0.09*	0.3	0.027		
s_2	1.0	*0.0*	0.6	0.0	**0.513**	
s_3	0.8	*0.81*	0.6	0.486		
s_1	0.3	*0.64*	0.3	0.192		
s_2	1.0	*0.08*	1.0	0.08	0.336	
s_3	0.8	*0.08*	0.8	0.064		
s_1	0.3	*0.08*	0.3	0.024		
s_2	1.0	*0.72*	0.6	0.432	**0.456**	
s_3	0.8	*0.0*	0.6	0.0		
s_1	0.3	*0.0*	0.3	0.0		
s_2	1.0	*0.01*	1.0	0.01	**0.082**	
s_3	0.8	*0.09*	0.8	0.072		
s_1	0.3	*0.08*	0.3	0.024		**0.595**
s_2	1.0	*0.01*	0.6	0.006	0.036	
s_3	0.8	*0.01*	0.6	0.006		
s_1	0.3	*0.08*	0.3	0.024		
s_2	1.0	*0.01*	1.0	0.01	0.042	
s_3	0.8	*0.01*	0.8	0.008		
s_1	0.3	*0.01*	0.3	0.003		
s_2	1.0	*0.0*	0.6	0.0	**0.057**	
s_3	0.8	*0.09*	0.6	0.054		

$V^0(x_0)$ が成立する。すなわち

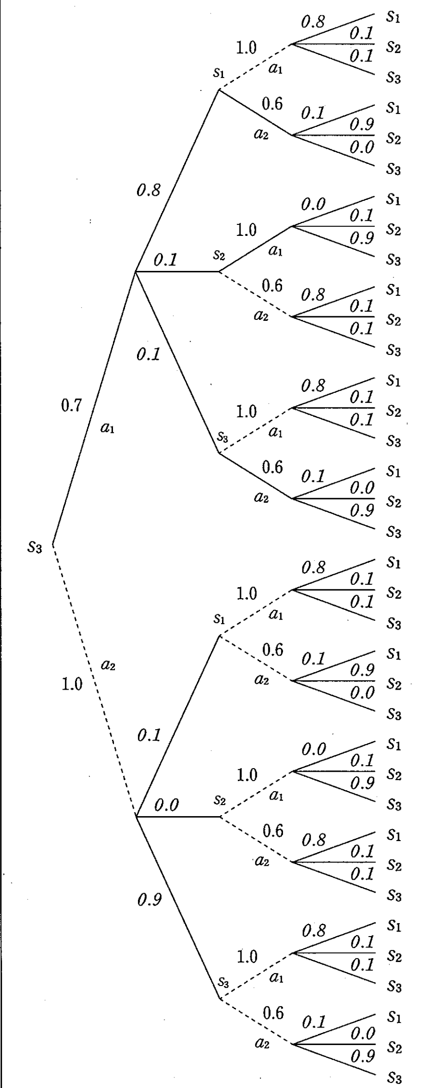
$$v^2(s_1)=V^2(s_1)=0.3,\qquad v^2(s_2)=V^2(s_2)=1.0,\qquad v^2(s_3)=V^2(s_3)=0.8$$
$$v^1(s_1)=V^1(s_1)=0.57,\qquad v^1(s_2)=V^1(s_2)=0.82,\qquad v^1(s_3)=V^1(s_3)=0.57$$
$$v^0(s_1)=V^0(s_1)=0.795,\qquad v^0(s_2)=V^0(s_2)=0.595,\qquad v^0(s_3)=V^0(s_3)=0.583 \tag{72}$$

である。

なお，図 2，3，4 では以下の略号を用いている：

$$履歴 = x_0\ \ r_0(u_0)/u_0\ \ p(x_1|x_0,u_0)x_1\ \ r_1(u_1)/u_1\ \ p(x_2|x_1,u_1)\ \ x_2$$

図 4　2 段過程に対する状態 s_3 からの全行動と最大分枝の選択



履　歴	終端	経路	最小	積	部分期	全期待
	0.3	0.64	0.3	0.192		
	1.0	0.08	0.7	0.056	0.304	
	0.8	0.08	0.7	0.056		
	0.3	0.08	0.3	0.024		
	1.0	0.72	0.6	0.432	**0.456**	
	0.8	0.0	0.6	0.0		
	0.3	0.0	0.3	0.0		
	1.0	0.01	0.7	0.007	**0.070**	
	0.8	0.09	0.7	0.063		**0.583**
	0.3	0.08	0.3	0.024		
	1.0	0.01	0.6	0.006	0.036	
	0.8	0.01	0.6	0.006		
	0.3	0.08	0.3	0.024		
	1.0	0.01	0.7	0.007	0.038	
	0.8	0.01	0.7	0.007		
	0.3	0.01	0.3	0.003		
	1.0	0.0	0.6	0.0	**0.057**	
	0.8	0.09	0.6	0.054		
	0.3	0.08	0.3	0.024		
	1.0	0.01	1.0	0.01	0.042	
	0.8	0.01	0.8	0.008		
	0.3	0.01	0.3	0.003		
	1.0	0.09	0.6	0.054	**0.057**	
	0.8	0.0	0.6	0.0		
	0.3	0.0	0.3	0.0		
	1.0	0.0	1.0	0.0	**0.0**	
	0.8	0.0	0.8	0.0		0.570
	0.3	0.0	0.3	**0.0**		
	1.0	0.0	0.6	0.0	**0.0**	
	0.8	0.0	0.6	0.0		
	0.3	0.72	0.3	0.216		
	1.0	0.09	1.0	0.09	0.378	
	0.8	0.09	0.8	0.072		
	0.3	0.09	0.3	0.027		
	1.0	0.0	0.6	0.0	**0.513**	
	0.8	0.81	0.6	0.486		

終端＝終端評価値＝$r_G(x_2)$

経路＝経路確率＝$p(x_1|x_0,\ u_0)p(x_2|x_1,\ u_1)$

最小＝最小評価値＝$r_0(u_0)\wedge r_1(u_1)\wedge r_G(x_2)$

積＝経路×最小

部分期＝部分期待値，全期待＝全期待値

さて，式(30)で定義された1変数最大値数列$\{V^0,\ V^1,\ V^2\}$は再帰式

$$\begin{aligned} V^2(x) &= r_G(x) \\ V^1(x) &= \operatorname*{Max}_{u\in U}\Big[r_1(u)\wedge\sum_{y\in X}V^2(y)p(y|x,\ u)\Big] \\ V^0(x) &= \operatorname*{Max}_{u\in U}\Big[r_0(u)\wedge\sum_{y\in X}V^1(y)p(y|x,\ u)\Big] \end{aligned} \tag{73}$$

を満たさないこと(脚注3)を確かめよう。そのために，新たに再帰式([4])

$$\begin{aligned} W^2(x) &= r_G(x) \\ W^1(x) &= \operatorname*{Max}_{u\in U}\Big[r_1(u)\wedge\sum_{y\in X}W^2(y)p(y|x,\ u)\Big] \\ W^0(x) &= \operatorname*{Max}_{u\in U}\Big[r_0(u)\wedge\sum_{y\in X}W^1(y)p(y|x,\ u)\Big] \end{aligned} \tag{74}$$

で定義される関数列$\{W^0,\ W^1,\ W^2\}$が$\{V^0,\ V^1,\ V^2\}$と異なることをみればよい。事実，冒頭のデータ(63),(64),(65)を用いて式(74)によって$\{W^0,\ W^1,\ W^2\}$を計算すると，

$$\begin{aligned} &W^2(s_1)=0.3, && W^2(s_2)=1.0, && W^2(s_3)=0.8 \\ &W^1(s_1)=0.6, && W^1(s_2)=0.82, && W^1(s_3)=0.6 \\ &W^0(s_1)=0.798, && W^0(s_2)=0.622, && W^0(s_3)=0.622 \end{aligned} \tag{75}$$

になる。式(72)の値と比較すれば，$(W^0,\ W^1,\ W^2)$は$(V^0,\ V^1,\ V^2)$に一致していないことがわかる。したがって，(脚注3)の反例が与えられたことになる。

さらに，式(14)で定義された1変数関数列$\{v^0,\ v^1,\ v^2\}$に対しては式(72)より

$$(v^0,\ v^1,\ v^2)=(V^0,\ V^1,\ V^2)$$

である。したがって，$(W^0,\ W^1,\ W^2)$は$(v^0,\ v^1,\ v^2)$にも一致していないから，上述の反例は同時に(脚注2)の反例にもなっている([4],[5],[11],[14],[15],[18]も参照)。

参考文献

[1] J.F. Baldwin and B.W. Pilsworth, Dynamic programming for fuzzy systems with fuzzy environment, Journal of Mathematical Analysis and Applications **85** (1982), 1-23.

[2] R.E. Bellman, Dynamic Programming, Princeton Univ. Press, NJ, 1957.

[3] R.E. Bellman and E.D. Denman, Invariant Imbedding, Lect. Notes in Operation Research and Mathematical Systems **52**, Springer-Verlag, Berlin, 1971.

[4] R.E. Bellman and L.A. Zadeh, Decision-making in a fuzzy environment, Management Science **17** (1970), B 141-B164.

[5] A.O. Esogbue and R.E. Bellman, Fuzzy dynamic programming and its extensions, TIMS/Studies in the Management Sciences **20** (1984), 147-167.

[6] R.A. Howard, Dynamic Programming and Markov Processes, MIT Press, Cambridge, Mass, 1960.
[7] 岩本誠一，動的計画の理論と応用，数学 **31** 巻4号，1979年，331-348.
[8] 岩本誠一，「動的計画論」，九大出版会，1987.
[9] S. Iwamoto, Associative dynamic programs, Journal of Mathematical Analysis and Applications, to appear.
[10] S. Iwamoto and T. Fujita, Stochastic decision-making in a fuzzy environment, J. Operations Res. Soc. Japan **38** (1995), No. 4, 467-482
[11] J. Kacprzyk, Decision-making in a fuzzy environment with fuzzy termination time, Fuzzy Sets and Systems **1** (1978), 169-179.
[12] R.E. Lucas Jr., Models of Business Cycles, Basil Blackwell, 1987; 清水啓典訳「マクロ経済学のフロンティア—景気循環の諸モデル—」，東洋経済新聞社，1988年.
[13] E.S. Lee, Quasilinearization and Invariant Imbedding, Academic Press, New York, 1968.
[14] 水本雅晴，「ファジィ理論とその応用」，サイエンス社，1988.
[15] 小田中敏男，「確率制御過程」，森北出版，1976.
[16] M.L. Puterman, Markov Decision Processes : discrete stochastic dynamic programming, Wiley & Sons. New York, 1994.
[17] M. Sniedovich, Dynamic Programming, Marcel Dekker, Inc. NY, 1992.
[18] W.E. Stein, Optimal stopping in a fuzzy environment, Fuzzy Sets and Systems. **3** (1980), 253-259.
[19] N.L. Stokey and R.E. Lucas Jr., Recursive Methods in Economic Dynamics, Harvard Univ. Press, Cambridge, MA, 1989.