1999年度日本オペレーションズ・リサーチ学会
秋季研究発表会

多段確率決定樹表について

01003676　九州大学　岩本誠一　IWAMOTO Seiichi

1　はじめに

動的計画法は、多変数実数値関数の**同時最適化**を１変数ずつ逐次最適化する方法である。これは、目的関数の**再帰性** recursiveness（**可分性** separability ともいう）と**単調性** monotonicity の下で可能である。このとき、**逐次最適化**の最適解が本来の同時最適解でなければならない。すなわち、逐次最適化よる最適値が本来の最適値に一致し、逐次最適点が同時最適点（全体の少なくとも一部を）を与えることである。動的計画法が最も適用される目的関数は**加法型関数**であり、**確定性の下**である。この状況では再帰性と狭義単調性は満たされている。したがって、ただちに動的計画法を適用することができる。実際、**最適性の原理**より再帰式を導いて、これを解いて本来の多変数同時最適化の最適解を求めることができる。通常、動的計画法とはこのことを指している。

2　非加法型評価

最近、不確実性下における経済動学やファジィ環境下における多段意思決定には**非加法型評価**基準が用いられている。しかし、ここでは加法型と同様なアプローチがとられている。

不確実性の下では非加法型関数の期待値を決定関数列（政策）について適当な政策空間上で最適化することになる。ここで動的計画法を適用しようとすると、再帰性と単調性の成立を確認する必要がある。これには、(1) 期待値は政策の関数であり、したがって、(2) 再帰性と単調性は**決定関数列に関する**性質であること、に注意する必要がある。このとき、単調性については政策空間の選択いかんに関わらず、本来の（期待値をとる前の）単調性より常に成り立つ。しかし、再帰性に関しては常に成り立つとは限らない。この議論は、「動的計画法とは何か?」、「状態とは何か？」に還元され、不確実性の下で**終端状態モデル論**を展開する必要がある。

3　繰り返し法対直接法

本報告では、（非加法型関数とは限らず）任意の評価関数を不確実性の下で逐次最適化する方法を二つ、（１）繰り返し法と（２）直接法、提示する。前述のように同時最適解を与えるという意味で、動的計画法が正しく機能しているかをチェックする簡単な方法として**多段確率決定樹表** multi-stage stochastic decision tree-table による方法が有効であることを示す。これは決定樹と決定表からなり、問題記述から最適解に至るまでが再帰式を解く順に構成されている。再帰式と樹表は１対１に対応しているので、繰り返し法の樹表と直接法の樹表が共通の最適解を与えていることが分かる。

4　評価系対政策空間

評価系は（１）単一と（２）複合に分かれる。前者には、加法型、乗法型、最大型（最小型）、終端型があり、後者には範囲、比、分散などが考えられる。政策には（１）原始、（２）一般、（３）マルコフ、の３種類があり、それぞれ対応する政策空間上での最適化が考えられる。次ページでは、一例として加法型評価系の期待値最適化問題

$$\begin{aligned}&\text{Max } E^{\mu}_{x_1}[r_0 + r_1 + r_G]\\ &\text{s.t.}\quad x_{n+1} \sim p(\cdot\,|\,x_n, u_n),\quad u_n \in U\end{aligned}$$

を繰り返し法による多段確率決定樹表で解いている。加法型評価系に対する最適政策は特にマルコフ政策になることが示される。また、直接法での樹表（省略）でも同一の最適解が得られることが分かる。

一般に、最適政策が一般政策空間内に、直接法と繰り返し法の２つの樹表によって求められる。また、樹表はファジィ決定過程や事前・事後の両条件付き決定過程に対しても利用できることを示す。

$$V^0(s_1) = \underset{\mu_0}{\mathrm{Max}}\,\underset{\mu_1}{\mathrm{Max}} \sum\sum_{(x_1,x_2)\in X\times X} [\,r_0(u_0) + r_1(u_1) + r_G(x_2)\,]p(x_1|s_1,u_0)p(x_2|x_1,u_1)$$

図1：状態 s_1 からの2段確率決定樹表（繰り返し法）

r_0	p_0	履歴	r_1	p_1	r_G	加法	p_0p_1	積	部分期	全期待
0.7 a_1	0.8	s_1	1.0 a_1	0.8	s_1 0.3	2.0	0.64	1.28	1.696	2.248
				0.1	s_2 1.0	2.7	0.08	0.216		
				0.1	s_3 0.8	2.5	0.08	0.2		
			0.6 a_2	0.1	s_1 0.3	1.6	0.08	0.128	**1.784**	
				0.9	s_2 1.0	2.3	0.72	1.656		
				0.0	s_3 0.8	2.1	0.0	0		
	0.1	s_2	1.0 a_1	0.0	s_1 0.3	2.0	0.0	0	**0.252**	
				0.1	s_2 1.0	2.7	0.01	0.027		
				0.9	s_3 0.8	2.5	0.09	0.225		
			0.6 a_2	0.8	s_1 0.3	1.6	0.08	0.128	0.172	
				0.1	s_2 1.0	2.3	0.01	0.023		
				0.1	s_3 0.8	2.1	0.01	0.021		
	0.1	s_3	1.0 a_1	0.8	s_1 0.3	2.0	0.08	0.16	**0.212**	
				0.1	s_2 1.0	2.7	0.01	0.027		
				0.1	s_3 0.8	2.5	0.01	0.025		
			0.6 a_2	0.1	s_1 0.3	1.6	0.01	0.016	0.205	
				0.0	s_2 1.0	2.3	0.0	0		
				0.9	s_3 0.8	2.1	0.09	0.189		
1.0 a_2	0.1	s_1	1.0 a_1	0.8	s_1 0.3	2.3	0.08	0.184	0.242	**2.791**
				0.1	s_2 1.0	3.0	0.01	0.03		
				0.1	s_3 0.8	2.8	0.01	0.028		
			0.6 a_2	0.1	s_1 0.3	1.9	0.01	0.019	**0.253**	
				0.9	s_2 1.0	2.6	0.09	0.234		
				0.0	s_3 0.8	2.4	0.0	0		
	0.9	s_2	1.0 a_1	0.0	s_1 0.3	2.3	0.0	0	**2.538**	
				0.1	s_2 1.0	3.0	0.09	0.27		
				0.9	s_3 0.8	2.8	0.81	2.268		
			0.6 a_2	0.8	s_1 0.3	1.9	0.72	1.368	1.818	
				0.1	s_2 1.0	2.6	0.09	0.234		
				0.1	s_3 0.8	2.4	0.09	0.216		
	0.0	s_3	1.0 a_1	0.8	s_1 0.3	2.3	0.0	0	**0**	
				0.1	s_2 1.0	3.0	0.0	0		
				0.1	s_3 0.8	2.8	0.0	0		
			0.6 a_2	0.1	s_1 0.3	1.9	0.0	0	**0**	
				0.0	s_2 1.0	2.6	0.0	0		
				0.9	s_3 0.8	2.4	0.0	0		