文　献　抄　録

Desoer, C. A.; The Bang Bang Servo Problem Treated by Variational Techniques, *Information and Control, Vol.* 2 (1959), 333-348

線型微分方程式

$$\dot{y}=Ay+f(t),\quad y(0)=c \tag{1}$$

で記述される制御系がある．$f(t)$ が制御項，A は実数要素の $n\times n$ 行列で固有値はみな非負実部をもつとする．与えられた初期位置 $y(0)$ から出発して最小時間で $y(t)=0$ となるような制御を求めんとする．もし制御に

$$|f_i(t)|\leqq\gamma_i \qquad (i=1,\cdots,n) \tag{2}$$

という制約があれば

［定理］最適制御が存在して，それは bang-bang 型の制御：

$$|f_i(t)|\equiv\gamma_i \qquad (i=1,\cdots,n)$$

であり，各 $f_i(t)$ は高々 $(n-1)$ 回しか符号を変えない．

この有名な定理の Bellman-Glicksberg-Gross (1956)｛本誌，4巻1号，48参照｝の証明は topological であるので，変分法でやるとわかり易い．

A の固有根を $\lambda_1,\cdots,\lambda_n$，それぞれの固有列 vector を $x_1,\cdots\cdots,x_n$ とすると

$$A(x_1,\cdots\cdots,x_n)=(x_1,\cdots,x_n)$$

$$A(x_1,\cdots,x_n)=(x_1,\cdots,x_n)\begin{pmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_n\end{pmatrix}$$

変換 $y=(x_1,\cdots,x_n)u$ をやって（1）をかき直し，積分して．$u(t_0)=0$ を入れると，問題は

$$-u_i(0)=\int_0^{t_0}e^{-\lambda_i t}\sum_{j=1}^{n}\alpha_{ij}f_j(t)dt$$

$$(i=1,\cdots,n;\ \text{for some } f(t) \text{ with } (2))$$

のような最小の t_0 を求めたい（ここに $(x_1,\cdots,x_n)^{-1}=(\alpha_{ij})$）．ここまでは Bellman らと同じだが，ここから変分法を使う．

（2）より

$$f_i(t)=\gamma_i\sin\theta_i(t),\ -\frac{\pi}{2}\leqq\theta_i(t)\leqq\frac{\pi}{2}(i=1,\cdots,n)$$

とおいて変分問題：

$$J[\theta_1,\cdots,\theta_n]\equiv\int_0^{t_0}e^{-\lambda_1 t}\sum_{j=1}^{h}\alpha_{ij}\gamma_j\sin\theta_j(t)dt \to\max$$

subject to

$$\int_0^{t_0}e^{-\lambda_i t}\sum_{j=1}^{n}\alpha_{ij}\gamma_j\sin\theta_j(t)dt+u_i(0)=0$$

$$(i=2,\cdots,n)$$

を考える．この最大値はもちろん t_0 の函数だが，最大値が $-u_i(0)$ になるまで t_0 を動かしてゆくとそのときの t_0 および $\theta(t)$ が，我々の最適制御の解を構成する．

例題

$$\begin{pmatrix}\dot{x}_1\\ \dot{x}_2\end{pmatrix}=\begin{pmatrix}0 & 1\\ 0 & -1\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix}+\begin{pmatrix}0\\ f_2\end{pmatrix}$$

$$|f_2(t)|\leqq1$$

について計算してある．　　　　　　（坂口　実）

White, G. M.; Penny Matching Machines, *Information and Control, Vol.* 2 (1959), 349-363.

player(I とする）が indifferent opponent（II とする）を相手に 2×2 行列 game をやる．game といっても相手が indifferent opponent，すなわち自分の利益のための学習をしないから1人 game である．

支払行列

I＼II	おもて h	うら t
おもて H	G_{Hh}	G_{Ht}
うら T	G_{Th}	G_{Tt}

は I，II に既知で，II が fixed but unknown strategy をとる．I は学習して II の strategy を estimate しながら最適にふるまう．

(i)　II が Bernoulli strategy をとるとき．

II が一定確率 $P(h),P(t)$ でそれぞれ表・うらを出しているとする．I が混合戦略 $(P(H),\ P(T))$ をとるときの期待額 E は

$$E=(\gamma P(h)-\delta)P(H)+\{(G_{Th}-G_{Tt})P(h)+G_{Tt}\}$$

ただし

$\gamma=G_{Hh}+G_{Tt}-G_{Ht}-G_{Th}$(>0と仮定する)

$\delta=G_{Tt}-G_{Ht}$

となるから，Ｉの best strategyは，$P(h)$ の estimate $\hat{P}(h)$ によって

$$\hat{P}(h)\begin{Bmatrix}>\\<\end{Bmatrix}\delta/\gamma \text{ ならば}\begin{Bmatrix}\text{おもて}\\\text{うら}\end{Bmatrix}\text{を出す}$$

ことになる．$\hat{P}(h)$ の求め方は例えば次のようである：$P(h)$ の事前分布として (0, 1) での一様分布を想定すると，今まで r 回表・s 回うらをⅡが出したときの，$P(h)$ の事後分布は $B(r+1, s+1)$ 分布だから，事後平均値をもって

$$\hat{P}(h)=E\{P(h)|\text{Ⅱが } r \text{ 回表・} s \text{ 回うらを出した}\}=\frac{r+1}{r+s+2}$$

とする．

(ii) ⅡがMarkov strategy をとるとき．

Ⅱが 推移確率 $P_h(h)$ および $P_t(h)$ の stationary Markov chain によって，おもて・うらを出しているとする．定常な chain だから周辺確率，例えば $P(h)$ は $P(h)=P(h)P_h(h)+P(t)P_t(h)$ から求まる．Ｉが混合戦略($P(H)$, $P(T)$) をとるとき，Ⅱが前回におもてを出したという条件つきのＩの期待額 E_h は

$$E_h=(\gamma P_h(h)-\delta)P_h(H)+\{(G_{Th}-G_{Tt})P_h(h)+G_{Tt}\}$$

同様に E_t は上で $P_h(h)\to P_t(h)$,$p_h(H)\to P_t(H)$ とかえたものに等しい．そこでＩの best strategy は

	Ⅱの推移確率のestimate		Ｉの best strategy		Ｉの期待額 $E=P(h)E_h+P(t)E_t$
	$\hat{P}_h(h)$	$\hat{P}_t(h)$	$P_h(H)$	$P_t(H)$	
(Ｉ)	$>\delta/\gamma$	$>\delta/\gamma$	1	1	
(Ⅱ)	$>$	$<$	1	0	
(Ⅲ)	$<$	$>$	0	1	
(Ⅳ)	$<$	$<$	0	0	

のようになる．もしⅡが実際は Markov strategy をとっているのに，Ｉがそれを知らずに，Bernoulli strategy をとっているものと誤解すれば，Ｉがどれだけ損をするかを計算している．　　(坂口　実)

Wilson, K. V. and Bixenstine, V. E.;
Forms of Social Control in Two-Person, Two-Choise Games
Behavioral Science, Vol. 7(1962), 92-102

非零和２人 game の最も簡単な2×2型のものにおける戦略の研究である．基本的な次の４つの型があると著者らは言っている：

(i) AP 制御 (*absolute* control over *personal* gain).

自分がどれをとっても相手にとっては同じだから自分の得になる方をえらぶというやり方，例えば，$\begin{pmatrix}5,5 & 5,1\\1,5 & 1,1\end{pmatrix}$ (左[右]側の数の行列がＩ[Ⅱ]の所得行列において player Ｉ，Ⅱが何れも 1st choice をえらぶ．

(ii) AO 制御 (*absolute* control over *other's* gain).

どれをとっても自分にとっては同じだから，相手の損になる方をえらぶ．例えば$\begin{pmatrix}5,5 & 5,1\\5,1 & 1,1\end{pmatrix}$においてＩ，Ⅱ (何れも 2nd choice をえらぶ．

(iii) CP 制御 (*conditional*control over *personal* gain)

相手の手がわからざれば自分の得になる手きまらず．

(Ⅳ) CO 制御 (*conditional* control over *other's* gain)

相手の手がわからざれば相手を損させる手きまらず，例えば$\begin{pmatrix}5,2 & 1,2\\1,2 & 5,2\end{pmatrix}$においては，ＩはCP 制御，Ⅱは CO 制御をとるだろう．Ⅱ[Ｉ]が例えは 1st choice をとるとわかっていてはじめて，Ｉ[Ⅱ]は 1st [2nd] choice をとれるからである．CP または CO 制御のときは具体的には mixed strategy をとることになる．

非0和２人2×2 game における一般の戦略が，これら４つの基本型の組合わせであることを多数の例で説明している．　　(坂口　実)