〈綜合報告〉 Dynamic Programming—発展の10年

坂口 実*

§0. はしがき

周知のように dynamic programming(以下単に DP と記す)というのは，米国 RAND Corporation の R. Bellman 博士の創始した理論で，彼の DP についての最初の論文と思われるもの [1] が 1952 年に出ているから今年でちょうど 10 年になる．その間，主として Bellman とその協力者によってなされてきた理論および応用的研究はおびただしい数に上る．それらは Bellman が書いた DP についての 3 冊の書物 [2], [3], [4] と数冊の RAND Report, 例えば [5], の中に参考文献として網羅されているが，数え上げたら 400 篇にもなろう．これらの論文は，DP の考え方が経営・管理・工学・制御その他の極めて多方面の技術的分野に如何に有用かつ強力であるかを明示しているが，DP が理論としても応用としても一応の反省期に達したと思われる現在，これら既成の成果を集成整理して新しい発展方面を用意することがなされねばならないわけである．Bellman の近著 [4] はこの目的にも沿うものであろうが，以下にこの書物をも参考にしながら，DP の 10 年間の発展の概略と今後の研究の方向とを，OR 的応用の立場に重点を置きながらながめることにしたい．なお参考までに，DP 全般の概要をわかり易くまとめたすぐれた解説的小論文として Bellman [6], Bellman [7], Dreyfus [9] などがあることを附記しておく．

§1. DP の本質，その歴史

DP とは問題の一つのながめ方である．LP のようにある特定の方程式系で表現されて特定の算法があるといったものではない．この点は一面不便であるが，また真の OR 研究者には大いに研究心をそそる所でもあろう．

DP の要点は，多段決定過程の問題を極値演算子をもった函数方程式，あるいはその離散的接近である漸化関係式にはめこむ(いわゆる invariant imbedding)ことである．漸化関係式は数理物理学などで古くから使われて来たし，現今でも queuing theory では標準的な数学的道具として使われる．しかしそれは何れも，逐次的な現象を記述する手段となっているので decision-making のために使われているのではない．decision-making に漸化関係式が使われたのは恐らく 1940 年代にさかのぼるだろう．数理統計学者 A. Wald の有名な sequential analysis の業績

* 電気通信大学 昭和 37 年 10 月 16 日受理　　経営科学第 6 巻 2 号

は1945年である．1951年頃から Bellman は多段決定問題に函数方程式を導入することに興味をもち，このような接近法を dynamic programming と命名した．彼の初期の論文の多くは，この新しい極値演算子つき函数方程式の解の存在1意性に関するものである．

1955年から Bellman は S. Dreyfus と協同して DP の計算面に努力を集中して来た．函数方程式を解くのに必要な数値計算の量を減少するよう種々の工夫をなした．それらは近著[4]にくわしい．

このような努力により DP の重要性が各方面で認識されてくるにつれて，多数の優れた研究者が米国内各地の研究機関において DP に関心を示すようになり，ここ数年前から米国 OR 学会年会の研究発表では，DP に関連した研究論文が提出されていない panel session はみられないまでになっている．

§2. いろいろの型の問題の展望

(i) 多段配分過程

DP は多段決定過程の問題を解く一つの強力な道具であるが，(a)すべての多段決定過程の問題は DP により解ける，とか(b)すべての DP の問題は多段決定過程の問題である，とか考えてはいけない．すべての多段決定過程の問題は，DP により定式化できるかも知れないが解けるというわけではない．問題によってはその大きさと複雑さが DP を無力にする場合も多いのである．(b)については，一見多段的観点をもたない問題に DP が使われることがある．次のような問題はこのよい例である：

$$\begin{cases} \sum_{k=1}^{n} g_k(x_k) \to \max\{\min\} \\ \sum_{1}^{n} x_k \leqq \{\geqq 0\} a(>0) \ ; \qquad x_k \geqq 0 (k=1, \cdots, n) \end{cases} \tag{2.1}$$

(第2式以下は拘束条件を示す．以後もこの記法に従う) n 個の活動があって，x_k を第 k 番目 $(k=1, \cdots, n)$ の活動に投入される資金，$g_k(x_k)$ はそのときの収益{費用}と考えられる．このような問題は OR の分野には大変多い．函数 $g_k(x)$ が皆 concave あるいは皆 convex のときは具体的な解析解があるが，その他の場合には解析的な一般解は求められていない．$g_1(x)=\cdots=g_k(x)=g(x)$，かつ $g(x)$ が唯一つの変曲点をもつ convex-concave (あるいは concave-convex) 函数のときでさえもそれを求めることは容易でない．しかし(2.1)を解く一つの解法が，DP の考えで次のように与えられることは明らかである：(2.1)での所求の最大値を $f_n(a)$ とおけば漸化関係式

$$f_n(a)=\max_{0 \leqq x \leqq a}[g_n(x)+f_{n-1}(a-x)] \qquad (n=1, 2, \cdots ; f_0(a) \equiv 0)$$

が成立するから，これを $f_1(a)$ から逐次求めてゆけばよろしい．

これをやってみると直ぐに，DP の考えを最大(小)問題に利用することの効用が明らかになる．

すなわち(a)極値でなく最大(小)値が割合少ない手間で求まる，(b)拘束条件例えば $a_i \leqq x_i \leqq b_i$ などが附加されても，却ってそれは計算を簡単にする，(c)離散的変数の問題でも同一に扱える，(d) $g_k(x)$ に滑らかさがなくとも(例えば階段函数・折線函数など)よい，等々．

(2.1)の2方向への拡張：

$$\begin{cases} \sum_{j=1}^{n} g_j(x_j) \to \max \\ \sum_{j=1}^{n} b_{ij}(x_j) \leqq a_i \qquad (i=1, \cdots, m) \\ x_j \geqq 0 \qquad (j=1, \cdots, n) \end{cases} \tag{2.2}$$

$$\begin{cases} \sum_{k=1}^{n} g_k(x_k, y_k) \to \max \\ \sum_{1}^{n} x_k \leqq a, \qquad x_k \geqq 0 \qquad (k=1, \cdots, n) \\ \sum_{1}^{n} y_k \leqq b, \qquad y_k \geqq 0 \qquad (k=1, \cdots, n) \end{cases} \tag{2.3}$$

は多次元の函数方程式を解かねばならぬという事態を生む．例えば(2.2)で $m=2, b_{ij}(x) \equiv b_{ij}x$ $(b_{ij}>0)$ とし，所求の最大値を $f_n(a_1, a_2)$ とかくと

$$f_n(a_1, a_2) = \max_{0 \leqq x_n \leqq \min_{i=1,2} \frac{a_i}{b_{in}}} [g_n(x_n) + f_{n-1}(a_1 - b_{1n}x_n, a_2 - b_{2n}x_n)] \qquad (n=1, 2, \cdots; f_0(a_1, a_2) \equiv 0) \tag{2.4}$$

を得るが，これを解くのに古典的な Lagrange 乗数法を併用すると計算量が相当軽減されることが知られている．すなわち固定した各 λ について

$$\begin{cases} \sum_{j=1}^{n} g_j(x_j) - \lambda \sum_{j=1}^{n} b_{1j}x_j \to \max \\ \sum_{j=1}^{n} b_{2j}x_j \leqq a_2, \qquad x_j \geqq 0 \, (j=1, \cdots, n) \end{cases}$$

を解き，その解を $(\bar{x}_1(\lambda), \cdots, \bar{x}_n(\lambda))$ として

$$\sum_{j=1}^{n} b_{1j}\bar{x}_j(\lambda) = a_1$$

をみたすまで λ を動かす(都合よいことに左辺が λ につき単調となる)と(2.4)の最適解が得られる [4][10]．(2.2), (2.3)のような定式化ができる OR の問題は，例えば多段装置の信頼性の問題，Knapsack の問題，Hitchcock-Koopmans の輸送問題などであるが [4] には例題について詳しい計算例がある．

なお [4：付録4] には問題(2.1)の形式的な解が次のように求められている．2つの函数 g と h との convolution を

$$g \otimes h(x) \equiv \max_{0 \leqq y \leqq x} [g(y)h(x-y)]$$

で定義すると，問題

$$\begin{cases} \prod_{k=1}^{n} \exp\{g_k(x_k)\} \to \max \\ \sum_{1}^{n} x_k = a(>0)\ ; \qquad x_k \geqq 0\,(k=1,\cdots,n) \end{cases}$$

の最大値は $e^{g_1}\otimes\cdots\otimes e^{g_n}(a)$ に等しい．convolution に対し容易に

(a) '擬 Lapace 変換' に対し

$$\max_{x\geqq 0}[e^{-zx}g\otimes h(x)] = \max_{y\geqq 0}[e^{-zy}g(y)]\cdot\max_{w\geqq 0}[e^{-zw}h(w)],$$

(b) 反転公式：$H(z)\equiv\max_{x\geqq 0}[e^{-zx}h(x)]$ とおくと

$$h(x)=\min_{z\geqq 0}[e^{zx}H(z)]$$

の成立がわかるから，この最大値の擬 Laplace 変換をとりそれを反転すると

$$e^{g_1}\otimes\cdots\otimes e^{g_n}(a)=\min_{z\geqq 0}\left[e^{za}\prod_{k=1}^{n}\max_{x\geqq 0}\exp\{-zx+g_k(x)\}\right]$$

となる．これの対数をとったものが問題(2.1)の所求の最大値であるというのである．

Bellman の DP 研究の初期には

$$f(x)=\sup_{0\leqq y\leqq x}[g(y)+h(x-y)+f(ay+b(x-y))] \quad (x\geqq 0\,;\ f(0)=0) \tag{2.5}$$

(ただし $g(x), h(x)$ は $g(0)=h(0)=0$ なる連続函数で，$0\leqq a, b<1$) の形の函数方程式の解の存在1意性に関する研究が多くあり，それらは彼の著書 [2] の第1章に再録されている．これは資金 x を2種類の事業に配分投資することを無限回続けてゆく ($1-a, 1-b$ が各事業の資本消耗率) といういわば仮想的な model を定式化したものだが，ここでもまた $g(x), h(x)$ が共に convex, あるいは共に concave なときは，最適政策は容易に求まりそれは簡単な構造をしている(解函数 $f(x)$ の形は必ずしも簡単でない)が，そうでないときは大変複雑であるという事情をみる．例えば変曲点のあるときに奇妙に不連続な最適政策が発生する例を挙げてある．

(2.5)で特に $h(x)\equiv 0$, $a=0$, とおくと問題(2.1)と関連があることは明らかであろう．また $g(x), h(x)$ が共にconvex で(2.5)が1意に解ける条件がある場合，(2.5)は新しい函数方程式

$$f(x)=\max[g(x)+f(ax),\ h(x)+f(bx)]$$

に転換される{逐次近似

$$f_n(x)=\max_{0\leqq y\leqq x}[g(y)+h(x-y)+f_{n-1}(ay+b(x-y))] \tag{2.6}$$

($n=1,2\cdots$; $f_0(x)\equiv 0$) を考えると帰納法で $f_n(x)=\max[g(x)+f_{n-1}(ax), h(x)+f_{n-1}(bx)]$, 極限にゆけばよし}．これは多段選択過程(後出)の問題に対応する式であって，特殊な場合でさえこれを解くのは大変難かしいが，その一例が [2] に出ている．

(ii) smoothing および在庫管理の問題

ある特定の状態で系を動かすのが望ましくて，この状態からの偏位の程度によって損失が生じまた系の状態を変化させるのにも費用がかかるとする．この系の操作全体からえられる効用を最

大ならしめるためには，一方の費用を他方の費用と釣り合わせながら中道をゆくのが最も有利であるということは OR の問題に広くみられる事態である．その典型的な例として平滑(smoothing)の問題と最適在庫管理の問題がある．

ある経済系の時刻 j における需要が r_j，実際の活動水準が x_j とする．すなわち $x_j \geqq r_j (j=1, 2, \cdots)$ であらねばならぬ．x_j-r_j および活動水準の変動 x_j-x_{j-1} に対しそれぞれ $g_j(x_j-r_j)$，$h_j(x_j-x_{j-1})$ の費用がかかるとすると，最小問題

$$\begin{cases} \sum_{j=1}^{n} \{g_j(x_j-r_j)+h_j(x_j-x_{j-1})\} \to \min \\ x_j \geqq r_j \qquad (j=1, 2, \cdots, n) \end{cases}$$

が生ずる．これを解くには

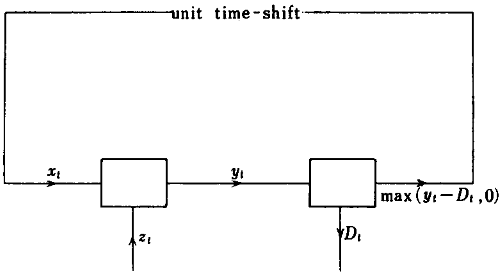
$$\begin{cases} \sum_{j=k}^{n} \{g_j(x_j-r_j)+h_j(x_j-x_{j-1})\} \to \min \\ x_j \geqq r_j \qquad (j=k, k+1, \cdots, n) \end{cases}$$

の最小値を $f_k(x_{k-1})$ とおくと

$$f_k(x_{k-1}) = \min_{x_k \geqq r_k} [g_k(x_k-r_k)+h_k(x_k-x_{k-1})+f_{k+1}(x_k)] \, (k=1, 2, \cdots, n\,;\ f_{n+1}(x) \equiv 0)$$

となるから，$f_1(x_0)$ を求めればよい．この問題のより進んだ優れた研究が例えば [11] にある．

在庫管理の問題で数学的 model を作り，それを dynamic programmig の方法で解析することは，Bellman その他による多くの研究 (Econometrica, Operations Research, Management Science などに散見) がある．在庫管理の問題は不確定の未来に直面して行動決定をなす



第 2-1 図

x_t: 第 t 期首における initial stock
y_t: 第 t 期首における starting stock
$z_t=y_t-x_t$: 第 t 期首における order の量 $(\geqq 0)$
D_t: 第 t 期中における需要量．一般に或る確率過程に従う．

問題の一つである．未来の不確定需要に見合う供給を stock するのに，供給過多や不足に対する損失についての仮定のもとで最適の stock 政策を求めようとする．ある種の簡単なしかし現実から余り遠くない仮定を設けた特殊の場合には，解が得られそれに経済的な意味がつけられる．

数学的 model は第 2.1 図に示すような一つの多段決定過程である．

確率的漸化式

$$x_{t+1} = \max(x_t+z_t-D_t, 0)$$

において，需要量 D_t が偶然量，注文量 $z_t=z_t(\cdots, D_{t-1};\, x_t)$ が制御項のとき損失函数の期待値

$$\sum_t a^t E[W_t(x_t, y_t;\, D_t)]$$

(有限又は無限の和，$0<a\leqq 1$ は割引率)を，或る適当な意味で最小にしようとする．最適決定 z_t，あるいは y_t は各 t に対し x_t の函数となることはいうまでもない．最適決定の時系列 $\left\{z_t\right\}_{t=1}^{\infty}$，$\left\{y_t\right\}_{t=1}^{\infty}$ をそれぞれ最適注文政策，最適在庫政策という．

いま最も簡単な場合を考えて，$\{D_t\}$ が同一分布に従う独立確率変数列で，密度函数 $\varphi(v)$ が既知とする．さらに W_t は皆同じ W とし

$$E\{W(x, y;\, D\}\equiv g(y-x)+\int_y^{\infty}\Psi(v, y)\varphi(v)dv$$

ただし，$g(z)$……量 z の注文に対する注文費用

$\Psi(v, y)$……在庫量 y のとき過大需要 $v(>y)$ に対する penalty cost

$\varphi(v)$……需要量分布の密度函数

とするとき

$u(x)$……initial stock x のとき最適注文政策を用いたときの total cost

とおくと

$$u(x)=\inf_{y\geqq x}\left[g(y-x)+a\left\{\int_y^{\infty}(\Psi(v, y)+u(0))\varphi(v)dv+\int_0^y u(y-v)\varphi(v)dv\right\}\right] \qquad (2.7)$$

の形の函数方程式が得られる [2].

ここで次の仮定がおかれていることに注意する：

(i) 政策決定者は注文をするかしないかの何れかで，返品は許されない．すなわち $y_t\geqq x_t$.

(ii) 注文がなされたときは，その全部の量が即座に引渡される．

(iii) 注文は期首に行なわれ，需要はその期の期末に集中して生ずると考える(penalty cost は次期費用に算入される)．

(iv) 需要量の従う確率分布法則 $\varphi(v)$ が既知であり，かつそれは無限遠の未来まで独立かつ同一である．

これらの仮定は，現実の事態に適応するためには，漸次取除きあるいは修正されねばならないことはもちろんで最近の多くの研究が

① 数学的理想化をできるだけ外して，実際の場合の設定に近づけること．

② 特に，需要量分布が partially known のときの数学的取扱い．

③ 既に得られている数学的な一般定理を，個々の場合に適用して，直ちに使用し得るような具体的な形にすること．

などに向けられている．主要な努力と結果を幾つか挙げると，(2.7)の形の方程式の解の存在性1意性などについての多くの研究(Bellman[2])，のほかに

(a) 注文費用・信用損失がいづれも比例的：

$$g(z)=kz\ ;\ \Psi(v,y)=p(v-y)\ (v\geqq y)$$

(k, p は正定数)とすると，$0<a<1$, $\varphi(v)>0$, $\int_0^\infty s\varphi(s)ds<\infty$ ならば，(2.7)の最適政策は定水準政策：

$$ap>k \text{ のとき } y^*(x)=\max(x,\bar{x})$$
$$ap\leqq k \text{ のとき } y^*(x)=x$$

となる．ここに $\bar{x}$ は

$$k=a\left(p\int_y^\infty \varphi(v)dv+k\int_0^y \varphi(v)dv\right)$$

の唯一つの根である [2].

(b) 上記の場合の少しの拡張として，注文に対する引渡しに何期かの遅れのある場合，$g(z)$ が凸函数の場合，2個以上の品種の在庫管理の場合が Bellman [2] にある．

(c) 注文費用・信用損失をそれぞれ一定値 K, P とする．さらに比例的な在庫保持費用(比例係数 h)をも考えると最適在庫方程式は

$$u(x)=\inf_{y\geqq x}\left[hy+K\mathrm{sgn}(y-x)+a\left\{(P+u(0))\int_y^\infty \varphi(v)dv+\int_0^y u(y-v)\varphi(v)dv\right\}\right]$$

となる．この最適政策はいわゆる s-S 注文政策，すなわちある適当な2正数 $0<s<S$ に対して

$$y^*(x)=S,\ (0\leqq x<s)\ ;\ =x,\ (x\geqq s)$$

となる(Arrow-Harris-Marschak [12])．この研究の当然の発展として最適注文政策が s-S 政策となるような model の特徴づけが研究されている [13].

(d) 需要分布が期ごとに変化する場合の研究．$u(x)$ を，initial stock が x で期ごとの需要密度函数が $\varphi_1, \varphi_2, \cdots$ のとき，最適注文政策を用いて望める総費用，とすると明らかに

$$u(x\ ;\ \varphi_1,\varphi_2,\cdots)=\inf_{y\geqq x}\Big[k(y-x)+L(y\ ;\ \varphi_1)$$
$$+a\left\{u(0\ ;\ \varphi_2,\varphi_3,\cdots)\int_y^\infty+\int_0^y u(y-v\ ;\ \varphi_2,\varphi_3,\cdots)\right\}\varphi_1(v)dv\Big],$$

ただし，r を単位量当りの売上利益として

$$L(y\ ;\ \varphi_1)\equiv\left(\int_0^y\{h(y-v)-rv\}+\int_y^\infty\{p(v-y)-ry\}\right)\varphi_1(v)dv.$$

Karlin [14] はこの解の研究をしている．例えば k, p, h が何れも比例的ならば $r>k$ のような当然な条件のもとに最適注文政策は定水準政策になる(しかし水準が期ごとに変る)．

(e) 需要分布に未知母数をふくむ場合の研究が Scarf [15] [16] にある．これは adaptive 制御の DP で解析されるがその解は普通2変数，すなわち initial stock x と未知母数の充足統計量 s とを含む．例えば需要分布の密度函数が指数型で

$$\varphi(v,\omega)=\beta(\omega)r(v)e^{-\omega v}\qquad (v>0, r(v)>0)$$

なことがわかっていて母数 ω の値だけが未知とすると，n 期目には ω に関する情報は充足統

計量

$$s_{n-1}\equiv(v_1+v_2+\cdots+v_{n-1})/(n-1)$$

につくされる．未知母数 ω に関する a priori 分布を想定すると例えば［定理］統計量の系列 $\{\bar{x}_n(s_{n-1})\}$ が存在して，第 n 期の在庫を $y_n^*=\max(\bar{x}_n(s_{n-1}),x)$ とするのが最適政策である．函数形 $\bar{x}_n(s)$ はもちろん想定した a priori 分布に依存してきまる．

（f） 需要が Poisson 過程で起るとすると continuous-time の過程が生ずる．Beckmann [17] は配達の遅れ時間も確率変数で，outstand order の経過時間 t のとき時間 Δt の間に配達される確率が $\mu(t)\Delta t+o(\Delta t)$ として，s-S 注文政策をとるときの過程の記述的性質を調べている．さらに，需要の起る時点のみでなく需要量も確率変数とした拡張もある．

およそ DP では，(2.7)型の方程式の解の存在定理や近似解の計算方式などは得られるが，解の定性的な性質を知ろうとすると多くの場合困難がともなう．そこで DP の他に非線型計画の手法や Markov 過程の理論を強力に使った接近法が，Stanford 大学の諸学者により進められた．それらの研究は論文集 [18] に集められているが，例えば，s-S 政策が最適になる在庫模型の特徴づけや配達に遅れのある場合の深い研究の他に，需要の確率分布法則が週期的に変動する場合の最適在庫政策，配達の遅れがあるとき s-S 政策をとった場合の安定作用特性等々の研究がある．

なおこの項全般，すなわち在庫管理の数学模型についての優れた解説的論文が Karlin [19] にあることを附記しておく．

(iii) 最適捜索(optimal search)

DP の数値解法がまた新しい DP の問題を生むことは面白い．例えば DP による定式化をやって

$$f_k(x)=\max_y G(x,y,f_{k-1}(y))$$

という形の方程式を生じたとき，これが解析的には解けなくて数値解法によらねばならぬとすると，右辺の max は enumeration によって計算するわけである．すなわち極めて多数の $y_1,\cdots,y_N$ をとり $G(x,y_i,f_{k-1}(y_i))\,(i=1\cdots,N)$ を計算してこれらの値を比較するのである．解の精度をある程度保証ししかも計算時間を短かくすますには．計算の各段階で能率的な政策選択を行なう必要があろう．この問題は捜索過程の重要さを示す一つの例として，大きくまた難しい問題であるが，以下に現在までに解かれている幾つかの面白い捜索問題を述べてみよう．

（a） 単峰函数(unimodal ft.)の最大点分離 [2] [4]――函数 $y=f(x)$ は $[0,\infty)$ で strictly unimodal とする．$[0,A]$ の中に高々 n 個の点 $x_1,\cdots,x_n$ を適当にとり値 $y_i=f(x_i)\,(i=1,\cdots,n)$ を計算することによって，$\max_{0\leq x\leq A}f(x)$ に到達する長さ 1 以下の部分区間が指定可能であるとき，$f(x)$ は $[0,A]$ で性質 S_n をもつという．$[0,A]$ で性質 S_n をもつような A の上限を F_n とすると，$\left\{F_n\right\}_0^\infty$ は Fibonacci 数列である．すなわち $F_n=F_{n-1}+F_{n-2}\,(n=2,3,\cdots;\ F_0=F_1=1)$．$F_{20}>10{,}000$ であるから，この捜索方法によると最大点の位置は，もとの区間の長さの 10^{-4} 以内の部分区間に高々 20 回の evalutaion により分離される．

（b） 凸函数の零点——$f(x)$ は連続かつ単調な凸函数で $f(0)=1, f(1)=-a(a>0)$ とする．$f(x)$ の零点を推定するのに N 回の evaluation の後に零点をふくむある区間を指定してその幅を min-max にせよ，という問題を Gross-Johonson [21] が解いた．[4] には詳しい数値例がある．

（c） 不良銅貨の問題 [22]——数学 puzzle の中で有名な問題である．N 個の同種の coin の中に高々2個の重いものが混っている．これらを最小回数の秤量で全部つきとめるにはどうしたらよいか？ [22] では DP によりこれを定式化し $N=80, 90$ のときの解が電子計算機により作表されている．

（d） 故障の最適点検 [23]——複雑な電子装置が N 個の成分より，各成分 r は n_r 個の部品より成る．どれかの成分の中のどれかの部品に故障がある．故障場所の先験確率，点検所要時間などを与えられて，最も早く故障場所をつきとめる捜索手法を求めよ．

その他興味ある例が文献に多数みられる．[24] [25] [26] 等々．

(iv) 多段選択過程

既に(i)でみられたように，多段配分過程といい選択過程といってもそれは厳然たる区別ではなくて，ちょっとした修正で互いに移り変るのである．Bellman は DP 研究の初期に goldmining の問題をやっている：埋蔵量それぞれ x, y の金鉱 A, B がある．ある採掘機械を A 鉱に用いると確率 p_1 で埋蔵量の $100r_1\%$ を掘り出すことができるが，確率 $1-p_1$ でこわれてしまって全然掘り出せない．この機械を B 鉱に用いると，同じように確率 p_2 で埋蔵量の $100r_2\%$ を掘り出すが，確率 $1-p_2$ でこわれてしまう．この機械を次々に A または B に使って，機械がこわれない間に最大量を掘り出すにはどうすればよいか？

求める最大量を $f(x,y)$ とおけば

$$f(x,y)=\max\begin{bmatrix}A: p_1\{r_1x+f((1-r_1)x, y)\}\\ B: p_2\{r_2x+f(x,(1-r_2)y)\}\end{bmatrix} \tag{2.8}$$

$(x, y\geqq 0; f(0,0)=0)$ の成立は容易であろう．Bellman によると，$0\leqq p_1, p_2<1$ ならば連続な解が1意的に存在して，最適政策が

$$\frac{p_1r_1x}{1-p_1}\begin{Bmatrix}>\\<\end{Bmatrix}\frac{p_2r_2y}{1-p_2} \text{ ならば } \begin{Bmatrix}A\\B\end{Bmatrix} \text{ をえらべ}$$

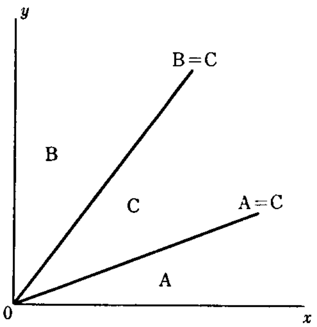
となる．

一般に方程式が

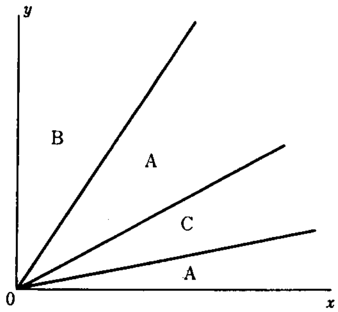
$$f(x)=\max\begin{bmatrix}A_1: & g_1(x)+f(T_1x)\\ A_2: & g_2(x)+f(T_2x)\\ \vdots & \vdots\\ A_k: & g_k(x)+f(T_kx)\end{bmatrix}$$

(T_1 等は与えられた点変換で shrinking($\|Tx\|<\|x\|$) な性質をもつ)の形となる問題は実際面には極めて多いのだが，これを解くことは $k=2$ の場合でさえ大変難かしい．例えば(2.8)でも A と B の中間の第3の選択 C: $p_3\{r_3x+r_4y+f((1-r_3)x, (1-r_4)y)\}$ が入ると，最適政策は優先決定

領域が予想されるように第 2.2 図(a)とならず図(b)のようになることもあり，最適政策の性質は殆んど何もわかっていない [2].



第 2-2 図(a)



第 2-2 図(b)

この型の問題でかなりの結果が得られつつある方面は例えば，幾つかの部品を幾つかの機械に次々にかける順序づけ(sheduling)の問題 [2], [4], shuttle process の問題 [27] や特に optimal stopping の問題 [28] [29] [30] であろう. 最後のものの1例だけをつぎに挙げよう.

結婚の問題 [29]: n 人の女性がいて at random に1人づつ貴君の前に現れる. 各 r 番目の女性に直面して貴君は彼女と結婚するか，または(もっとよい女性が後に現れることを期待して)彼女を流して次の$(r+1)$番目の女性に対面するか，どちらかに決めねばならぬ. 各女性のよさに順位がつけられて(順位1が最良)，順位 i の女性と結婚できることの効用を U_i とする. 貴君の効用の期待値を最大にするにはどうすればよいか？

$U(s,r)$…r 番目の女性がみかけの順位 s(今までに対面した r 人の中では順位 s ということ)のとき，以後最適政策で得られる効用

とすると

$$U(s,r)=\max\begin{bmatrix} S:\ \displaystyle\sum_{i=s}^{s+n-r}\binom{i-1}{s-1}\binom{n-i}{r-s}U_i\Big/\binom{n}{r}^{-1} \\ C:\ \displaystyle\sum_{s'=1}^{r+1}U(s',r+1)/(r+1) \end{bmatrix},$$

右辺 [ ] 内の上式の summand は $U_i P_r\{$彼女の真の順位が $i \mid r$ 番目の女性がみかけの順位 $s\}$ である. [定理] $U_1=1, U_2=\cdots=U_n=0$ とすると最適政策は n が非常に大きいとき漸近的に,「およそ $ne^{-1}\fallingdotseq 0.368\,n$ 人の女性をみた後，始めてみかけの順位1になった女性と結婚せよ」となる.

(v) Markov 型決定過程

N 個の可能な状態をもつ Markov 連鎖があってその遷移確率行列を

$$P\doteq(p_{ij}\mid i=1,\cdots,N)$$

とする(p_{ij} は $j\to i$ の遷移確率である). 遷移が時刻 $n\Delta(n=0,1,2,\cdots)$で起り，時刻 t のときの状態確率 vector を $x(t)$とすると

$$x_i(t+\Delta)=\sum_{j=1}^{N} p_{ij}x_j(t), \qquad x_i(0)=c_i \qquad (i=1,\cdots,N).$$

行列記法でかくと

$$x(t+\Delta)=Px(t), \qquad x(0)=c.$$

いま遷移行列 P が助変数 q に依存するとし，状態確率 $x_1(t)$ が常に最大になるようにこれを制御するならば

$$\begin{cases} x_1(t)=\max\limits_{q}\sum\limits_{j=1}^{N} p_{1j}(q)x_j(t-\Delta) \\ x_i(t)=\sum\limits_{j=1}^{N} p_{ij}(q^*)x_j(t-\Delta), \quad (i=2,\cdots,N) \end{cases}$$

が成立する．ここに q^* は上式での最大に到達する q のことである．

ここで2方向の拡張がある．その一つは問題の連続型接近である．p_{ij} は負であってもよいとする．それには，$x_i(t)$ は時刻 t における財 i の現在量，p_{ij} は財 j の単位量当りの，財 i の $[t, t+\Delta]$ における生産または消費(負の場合)の量と考えればよい．つぎに $\Delta\to 0$ の極限を考えると

$$\begin{cases} \dfrac{dx_1}{dt}=\max\limits_{q}\sum\limits_{j=1}^{N} b_{1j}(q)x_j(t), \qquad x_1(0)=c_1 \\ \dfrac{dx_i}{dt}=\sum\limits_{j=1}^{N} b_{ij}(q^*)x_j(t), \qquad x_i(0)=c_i \qquad (i=2,\cdots,N) \end{cases} \tag{2.9}$$

が得られる($p_{ij}=b_{ij}\Delta$, $(i\neq j)$; $=1-b_{ii}\Delta$, $(i=j)$ とおけ)．

他の方向の拡張は過程

$$x(n+1)=\max_{q} P(q)x(n), \qquad x(0)=c(\geqq 0) \tag{2.10}$$

が考えられることである．ここに右辺の max は成分ごととする．すなわち $q=(q_1,\cdots,q_N)$ が vector 助変数で $[P(q)x(n)]_i=\sum_{j=1}^{N} p_{ij}(q)x_j(n)$ を最大にするには q_i だけ動かす，他の q_j とは interaction がないとするのである．Bellman [2] には方程式(2.9)をもっと一般にしたものの解の存在性1意性，(2.10)の過程の漸近的性質などについての研究がある．一端を挙げると［定理］$p_{ij}(q)$ が皆正ならば(a)正定数 λ が唯一つ存在して vector 方程式 $\lambda y=\max\limits_{q} P(q)y$ が正の解 $y>0$ をもつ．この解は正の定数因子を除いて1意で $\lambda=\max\limits_{q}(A(q)$ の最大絶対値の固有根)．(b) (2.10)の過程は漸近的に $x_i(n)\underset{(n\to\infty)}{\sim} k(c)\lambda^n y_i (i=1,\cdots,N)$ ここに $k(c)$ は定数因子，y および λ は(a)におけるものである．

報酬をともなう Markov 型決定過程が多くの OR 的応用をもつことを Howard [31] が示した．状態 j において決定 q をとるとき，immediate return $a_j(q)$ を生ずるとともに，遷移確率 $p_{ij}(q)$ $(i=1,\cdots,N)$ をもって状態 i に移るとする．状態空間から決定空間への函数 $q=f(j)$ が決定函数(decision ft.)で，決定函数の列 $\left\{f_n(j)\right\}_{n=1}^{\infty}$ が政策(policy)である．決定函数 $f(j)$ に対する immediate-return vector，遷移行列をそれぞれ

$$r(f)=\begin{bmatrix} a_1(f(1)) \\ a_2(f(2)) \\ \vdots \\ a_N(f(N)) \end{bmatrix}, \qquad P[f]=(p_{ij}(f(j))|i,j=1,\cdots,N)$$

とかくと，政策 $\{f_n\}$ を用うるときの total expected return は

$$V(f_1,f_2,\cdots)=r(f_1)+\sum_{n=1}^{\infty}\beta^n P[f_1]P[f_2]\cdots P[f_n]r(f_{n+1})$$
$$=r(f_1)+\beta P[f_1]V(f_2,f_3,\cdots) \tag{2.11}$$

となる．ここに $0<\beta<1$ は1期後の収入に対する割引率である．これを vector 最大にする政策 $\{f_n(\cdot)\}$ を DP により求めることを Howard の書物 [31] ではやっているが，最近 Blackwell [32] は彼の結果を数学的にもっと完全かつ見易いものにした．それによると例えば次のような結果がある：まず f，次に政策 $\pi=\{f_n\}$ で続ける政策を (f,π) とかく．$V(\pi_1)\geqq V(\pi_2)$ のことを $\pi_1 \supseteqq \pi_2$ とかく．すると

［定理］ (a)すべての f に対し $\pi^* \supseteqq (f,\pi^*)$ ならば，π^* が最適(すなわち，すべての π に対し $\pi^* \supseteqq \pi$)．(b)ある f に対し $(f,\pi)\supset\pi$ ならば $f^\infty\equiv\{f,f,\cdots\}\supset\pi$.

［定理］ $\{q|a_j(q)+\beta(p_{1j}(q),\cdots,p_{Nj}(q))V(f^\infty)>V(f^\infty)$ の第 j 成分$\}$ が(a)すべての $j=1,\cdots,N$ に対し空集合ならば f^∞ が最適，(b)ある j に対し空集合でなければ，その他の全部の j に対し $g(j)=f(j)$ とおくと．$g^\infty\supset f^\infty$.

これが Howard のいわゆる policy-improvement-routine である．系として

［定理］ 決定空間が有限ならば，stationary(すなわちある f に対し $\pi=f^\infty$)な最適政策が存在する．

最適政策 $\pi^*=f^\infty$ を用うるときの total expected return は(2.11)より

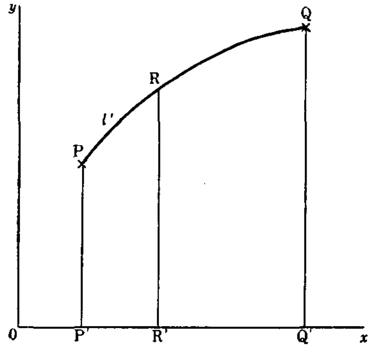
$$V(\pi^*)=(I-\beta P[f])^{-1}r(f)$$

となる．$\beta=1$ の場合は(2.11)が発散するおそれがあるので，Markov 連鎖の有名な漸近的性質を使ってもっと精密にやらねばならない．

報酬をともなう Markov 型決定過程の具体的応用例として，例えば [31] には taxi-cab の問題，[4] にはゴムタイヤ製造の問題，自動車の取替問題の詳しい数値例があり，また Herniter-Magee [33] には商店の広告に適用した例がある．

§3. 変分法の新しい形成

古典的な変分法は，函数を函数空間内の一点のように考えて極値を興えるものを求めるのであったが，DP による新しい考え方は，変分問題の解函数を一つの連続型多段決定過程の連続な最適政策として考えこれを求めようというのである．例で説明しよう(第3.1図)．2定点 P, Q を結ぶ定長 l の曲線の中で縦線図形 PP′Q′Q の面積を最大にするものを求めよ(等周問題)．この解曲線上 P, Q の中間に点 R をとる．$\widehat{\mathrm{PR}}$ の長さを l' とすると，部分曲線 $\widehat{\mathrm{RQ}}$ は2定点 R, Q, 定長 $l-l'$ を与えられたときの等周問題の解曲線となる．これは principle of optimality のた



第 3-1 図

めに他ならない．ここで最適政策は，y を x で表わした式よりもむしろ，$\dfrac{dy}{dx}$ を y および $l-l'$ で表わした式により表わされる．およそ，多段決定過程の最適政策を求めるのには，ある固定された初期位置から続く全部の決定列を定めるよりも，過程の現在位置の状態により次位置に対する決定を絶えず定めてゆく方が，ずっと簡単で自然でもある．変分法のこの新しいやり方の一つの利点はここにある．この事実は幾何学的には，解曲線を点の軌跡と考える代りに接線の envelope と考えるのである．前者の global に対して後者は local という意味で両者は dual である．

変分の問題

$$\begin{cases} J[y] \equiv \displaystyle\int_0^T F(x, y)\,dt \to \max \\ \dfrac{dx}{dt} = G(x, y), \qquad x(0) = c \end{cases} \tag{3.1}$$

は別に事新しい問題ではなくて函数 y を消去すれば容易に古典的な形

$$\begin{cases} J[x] \equiv \displaystyle\int_0^T f(x, \dot{x})\,dt \to \max \\ x(0) = c \end{cases}$$

になる．所が不等式の拘束条件が附加されて，例えば

$$\begin{cases} J[y] \equiv \displaystyle\int_0^T F(x, y)\,dt \to \max \\ \dfrac{dx}{dt} = G(x, y), \qquad x(0) = c \\ 0 \leqq y(t) \leqq x(t), \qquad (0 \leqq t \leqq T) \end{cases} \tag{3.2}$$

のようになると古典的な解析方法では力が及ばなくなってくる．この形の変分問題は数学的に興味があるばかりでなく OR の分野で大変重要である．多くの多段決定問題は continuous analogue をとると(3.2)のような変分問題となる．例えば(2.6)の多段配分過程については

$$\begin{cases} J[y] = \displaystyle\int_0^\infty (g(y) + h(x-y))\,dt \to \max \\ \dfrac{dx}{dt} = -(ay + b(x-y)), \qquad x(0) = c \\ 0 \leqq y(t) \leqq x(t) \end{cases}$$

となる．(3.2)の所求の最大値を $f(c, T)$ とおくと，函数 F, G の適当な解析性が仮定されていれば，偏微分方程式

$$\frac{\partial f}{\partial T} = \max_{0 \leqq v \leqq c}\left[F(c, v) + G(c, v)\frac{\partial f}{\partial c}\right] \tag{3.3}$$

が成立するをみるのは容易である(最適性原理により $f(c,T)=\max_{0\leq v\leq c}[F(c,v)\Delta T+f(c+G(c,v)\Delta T, T-\Delta T)]$ だから). はじめの函数としての拘束条件 $0\leqq y(t)\leqq x(t)$ が転じて数としてのそれ $0\leqq v\leqq c$ となった. 最適政策 v^* は c, T 2変数の函数であるから解析が面倒になる. この種の方程式の一般的な解の技巧というものはないようである. しかしある種の問題は簡単な解の構造をもっている. 例えば問題(3.2)において特に $F(x,y)\equiv G(x,y)$ ならば(最終値制御), $\max_{0\leq y\leq x} G(x,y)=G(x,y^*(x))$ とおくと $y^*(x)$ が最適政策である. すなわち $x(T)\to\max$ が目標の最終値制御では, 瞬間変化率を常に最大に保つように制御すればよろしい.

問題(3.2)から方程式(3.3)を導いたように, 問題(3.1)の本質を invariant imbedding したものが方程式

$$\frac{\partial f}{\partial T}=\max_v\left[F(c,v)+G(c,v)\frac{\partial f}{\partial c}\right]$$

である. これから, 古典理論における(3.1)の Euler 方程式(Lagrangian の第1変分を0に等置したもの)が得られることは次のようである: 右辺で最大に到達する v を $v^*(\mathrm{c},T)$ とかけば

$$\begin{cases} f_T=F(c,v^*)+G(c,v^*)f_c \\ 0=F_v(c,v^*)+G_v(c,v^*)f_c \end{cases}$$

であるから f_c, f_T を解き f_{cT} と f_{Tc} とを等置すればよい. (3.1)以外の型のいろいろの古典的変分問題についても, 連続型多段決定過程として考えて得られる(3.3)のような新しい Euler 方程式がそれぞれ対応する [34]. Euler, Lagrange, Weierstrass ほかの数学者が開発した古典的変分法では, 数値的結果を得ることよりもむしろ解の存在と単一性の定理に重きがおかれた. しかしながら現今の高速計算機の出現は数学的思考に大きな影響を与えた. われわれの利用できるこの強力なる道具をもって, 解の elegance のみでなく計算実行の可能性をも検証できるのである. 解析学で高い価値をもつ Euler 方程式は, 今や数値解法の観点からも再評価されねばならない.

(i) 最適制御と最適軌道

連続型であっても, 離散型であっても制御過程を分類してみると

制御目標 / 制御時間	(a) overallの制御	(b) 最終値制御	(c) (a), (b)の組合せ
(1)制御時間が固定されているもの	(a_1)	(b_1)	(c_1)
(2)制御時間が可変のもの	(a_2)	(b_2)	(c_2)

のようになるであろう. 第(2)型は制御時間長 T があらかじめ与えられていないで, その替りに終結面 $\mathbf{C}$ が与えられていて, 軌道 $x(t)$ が初めて $\mathbf{C}$ に入ったときの時刻 τ_e(終結時刻という)において制御を終えるのである. 終結時刻 τ_e は, 制御者が間接に制御できる変数である. 不等式拘束条件を除いた最も簡単な形でこれらを例示すれば, 例えば $\dot{x}=G(x,y), x(0)=c$ という拘束条件の下に

$$(\mathrm{a}_1)\text{型}:\qquad J[y]\equiv\int_0^T F(x,y)\,dt\to\max \tag{3.1 再}$$

$$(\mathrm{b}_1)\text{型}:\qquad J[y]\equiv H(x(T),y(T))\to\max$$

$$(\mathrm{a}_2)\text{型}:\qquad J[y]\equiv\int_0^{\tau_e} F(x,y)\,dt\to\max \tag{3.4}$$

$$(\mathrm{b}_2)\text{型}:\qquad J[y]\equiv H(x(\tau_e),y(\tau_e))\to\max$$

$$(\mathrm{c}_2)\text{型}:\qquad J[y]\equiv\int_0^{\tau_e} F(x,y)\,dt+H(x(\tau_e),y(\tau_e))\to\max$$

などである．いわゆる最速制御は(3.4)で $F(x,y)=-1$ の場合である．

これらの型の中で，理論的には，最も本質的な型は(b_2)である．実際，変数の次元を1つ上げることによって，第(1)型は第(2)型に，(a)型は(b)型に転ずる．すなわち，第(1)型において新しい位置変数 $\xi(t)$ を導入して，拘束条件 $\dot{\xi}=-1$，$\xi(0)=T$ を附加し，終結面を $\mathbf{C}=\{(x,\xi)\mid\xi=0\}$ と定義すれば第(2)型となる．また(a_2)型において新変数$\zeta(t)$を導入して

$$\begin{cases} J[y]\equiv\zeta(\tau_e)\to\max \\ \dot{x}=G(x,y), \qquad x(0)=c \\ \dot{\zeta}=F(x,y), \qquad \zeta(0)=0 \end{cases}$$

と考えれば，これは(b_2)型に他ならない．

このことはしかし，(b_2)型の解法だけを研究すればよいということにはならない．およそ DP の解法では，変数の次元を1つでも上げることが計算実行の困難さを累乗的に増大させるからである．(3.3)のような 'Euler 方程式'（これを制御問題(3.2)の主要方程式とよぶ）を実際に(解析的に，あるいは数値的に)解いて最適軌道を求めることが航空工学などで行われているようである．例えば [4] には rocket の minimum-climb-time の問題，人工衛星の最適軌道問題，原子炉中毒を防止するための最適操業休止の問題などが詳しく解かれている．

(ii) bottleneck 型の制御問題

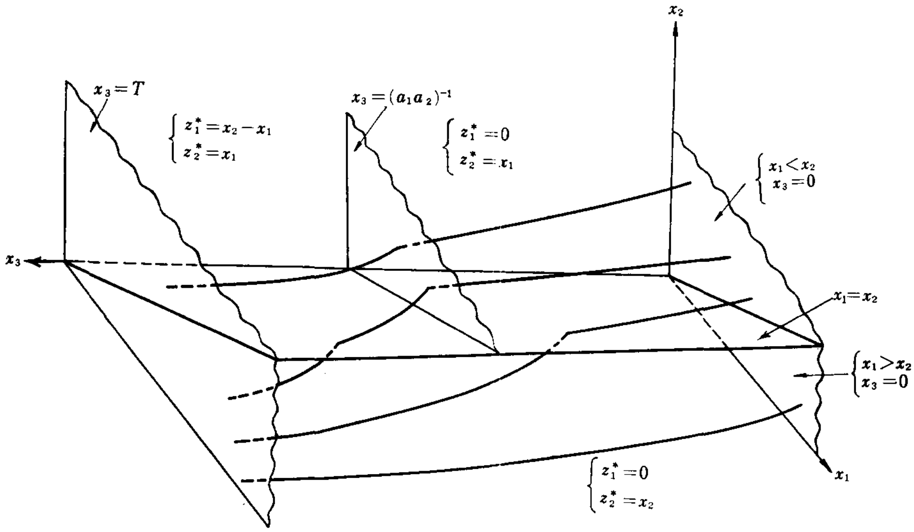
線型過程の(b_1)型制御で bottleneck 型の拘束条件があるもの：

$$\begin{cases} \alpha\cdot x(T)\to\max \\ \dot{x}=A_1x+A_2z, \qquad x(0)=c \\ B_1z\leqq B_2x, \qquad (0\leqq t\leqq T) \\ x,z\geqq 0 \end{cases} \tag{3.5}$$

のような問題は生産・経営・管理などの技術部門に極めて多い．例えば林学の分野，化学工業の多くの方面，多段構造になっている生産工業の各部門等々に現われるであろう．Bellman の書物 [2] にはその最も易しい例題として次のものが出ている：関連する2工業，例えば，自動車工業と製鉄工業とがある．時刻 t における自動車と鉄鋼の現在量をそれぞれ $x_1(t),x_2(t)$ とする．$x_2(t)$ を絶えず配分して自動車工業へ $z_1(t)$，製鉄工業へ $z_2(t)$ を投入する．一定期間 T 後の $x_2(T)$ を最大にするにはどのように配分してゆけばよいか？　定式化すると

$$
\begin{cases}
x_2(T)\to\max \\
\dot{x}=\begin{pmatrix} a_1 & 0 \\ -1 & a_2 \end{pmatrix} z, \qquad x(0)=c \\
\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} z \leqq x \\
x, z \geqq 0
\end{cases}
\tag{3.6}
$$

となる．$a_1, a_2(>0)$は与えられた転換係数である．

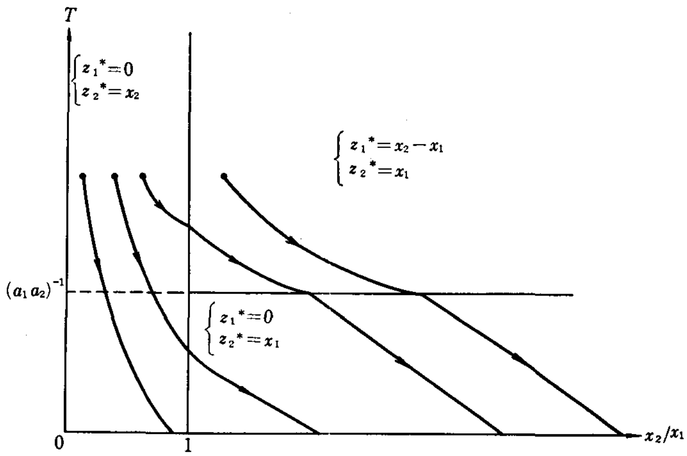


第 3-2 図

求める最大値を $f(c, T)$ とおくと，この制御過程の主要方程式は

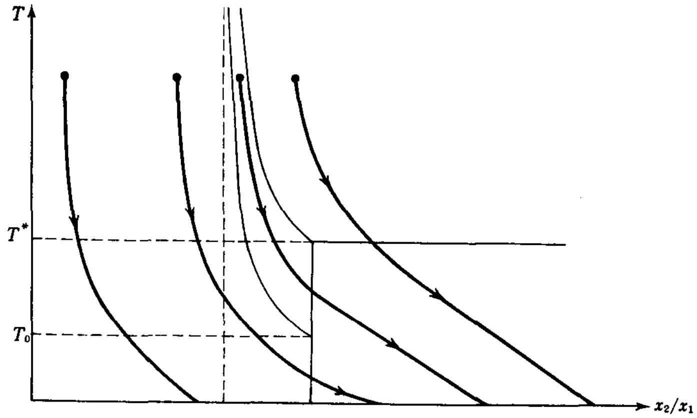
$$
\frac{\partial f}{\partial T}=\max_{\substack{z_1, z_2\geqq 0 \\ z_2\leqq c_1 \\ z_1+z_2\leqq c_2}}\left[\left(a_1\frac{\partial f}{\partial c_1}-\frac{\partial f}{\partial c_2}\right)z_1+a_2\frac{\partial f}{\partial c_2}z_2\right]
$$

となる．この右辺の最大問題が線型計画(LP)に他ならないことに注意すれば，この主要方程式は次のように解ける：最適制御および最適軌道は第3.2図，あるいは簡略に第3.3図で表わされ



第 3-3 図

る．最適制御の shifting boundary がいつも平面であるとは限らないのであって，例えば Bellman-Lehman [35] によると，(3.6)を少しく一般化した問題に対して最適制御が第 3.4 図のように表わされるのである．



第 3-4 図

また Bellman [2] によるとこの種の主要方程式の解法には，線型計画の双対定理から得られる知識を利用すべきである：問題(3.5)を直して

$$\begin{cases} \alpha\cdot x(T)\to\max \\ \dot{x}=Az, \qquad x(0)=c \\ Bz\leqq x \\ x,z\geqq 0 \end{cases}$$

とする．$x(t)=c+\int_0^t Az\,dt_1$ であるから，x を消去して

$$\begin{cases} \int_0^T \beta\cdot z\,dt\to\max \\ Bz-\int_0^t Az\,dt_1\leqq c \\ z\geqq 0 \end{cases}$$

(ここで $\beta=A'\alpha$ とおいた)の双対

$$\begin{cases} \int_0^T c\cdot w\,dt\to\min \\ B'w-\int_t^T A'w\,dt_1\geqq\beta \\ w\geqq 0 \end{cases}$$

を利用するのである(第 3.5 図)．

(iii) 線型過程の最適制御

制御問題では線型過程の解析が最も進んでいる．DP を用いる研究法は，1956 年 Bellman 等

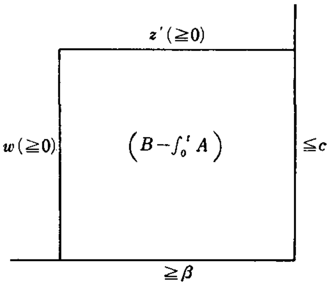
[36] がいわゆる bang-bang 原理を証明したことで大いに注目された．それは線型過程の最速制御

$$\begin{cases} J[y]\equiv\tau_e\rightarrow\min & (\text{ただし } \mathbf{C}=\{x=0\}) \\ \dot{x}=Ax+y, & x(0)=c \\ |y_i(t)|\leqq 1 & (i=1,\cdots,n) \end{cases}$$

(ここに A は実数要素の $n\times n$ 行列とする.) において，もしも A のすべての固有根が負の実部をもつならば，いわゆる bang-bang 型の

$$|y_i^*(t)|\equiv 1 \qquad (i=1,\cdots,n)$$

が最適制御である，というのである．



第 3-5 図

線型過程の中に加法的な random 攪乱が介入する場合の最適制御が最近盛んに研究されている．確率論的な困難をさけるために問題の離散型接近をとって例えば

$$\begin{cases} Q_N(v_0,\cdots,v_{N-1})\equiv u_N{}^2+\lambda\sum_{i=0}^{N-1}v_i{}^2\rightarrow\min \\ u_{k+1}=au_k+v_k+r_k, \qquad u_0=c \\ \{r_k\}\ \text{は確率法則既知の確率過程} \end{cases} \tag{3.7}$$

とするのである．これについては後述する．

(iv) 微分 game

制御者が複数で，しかも相反する目的をもって対抗する制御過程である．(b_2)型でいえば

$$\begin{cases} J[\phi,\psi]=H(x(\tau_e))\longrightarrow\max_{\phi}\text{-}\min_{\psi} \\ \dot{x}=G(x\,;\,\phi,\psi), \qquad x(0)=c \\ \text{終結面 } \mathbf{C} \end{cases}$$

のようになる．主要方程式は，game の値を $V(c)$ とおくと

$$\max_{\phi(0)}\min_{\psi(0)}\left(G(c\,;\,\phi(0),\psi(0))\cdot\frac{\partial V}{\partial c}\right)=0$$

となる(初めの Δt 時間に位置 $c'=c+G(c\,;\,\phi,\psi)\Delta t+o(\Delta t)$ まで移り $V(c)=\max_{\phi[0,\Delta t]}\min_{\psi[0,\Delta t]}V(c')$). この主要方程式の解析的な解法について Isaacs [37] のほか Berkovitz, Fleming, Scarf [38] などの研究がある．微分 game の具体的応用例については軍事的な model が多いが，Isaacs [37] の中にある多数の例のほかに例えば Zachrisson [39] の戦車の戦闘，Weiss [40] の Lanchestsr の戦闘 model についての論文がある．

§4. DP の今後の見通し

これからの DP はどのような方向にさらに発展してゆくであろうか？ 次の3方向があろうと思われる：(a)数値解法実行上の特殊な工夫の発展——最近の計算機械の進歩は，10 年以前には数学では扱えない分野と考えられていた多次元極値問題(§2(i)で述べたような)の解法に大い

に貢献している．簡単な問題(2.1)にしても一般解は現在のところまだ得られてないが，いつかは与えられるかも知れぬ．個々の問題について拘束条件の構造と系の発生的特徴とを利用したいろいろ特殊なくふうを考案すべきである．数理科学の実用的な発展のためにはこれは重要であろう．いままでに，Lagrange 乗数法を利用して解函数の次元を減らすくふう(§2(i))や，Legendre 多項式による函数近似を利用して数値計算量を軽減する考案 [4] [41] などがあるが，その他にも興味ある技法が [2] [4] の中の多数個所にみられる．

(b) feedback と adptive 制御——§3(iii)でもふれたように，線型過程の中に加法的な random 攪乱が介入する場合の最適制御が最近盛んに研究されている．さらにその攪乱の確率法則が partially known で何個かの未知母数を含むとき，それらを推定しながらそれに基づいて最適制御をしてゆくといういわゆる adaptive 制御の研究も，特に電気工学関係の研究者により進められている．このような不完全情報の問題は，その理論的解析は相当困難にはなるが興味ある多くの内容を含んでいる．過程が進行するにつれて制御機構に段々と情報が付加されて，それは自働的にあるいは入念な探査を経て蓄積される．制御機構は学習する(learn)，すなわち経験(観測と解析との適当な組合わせをいう)にもとづいてその特性(performance)を改良することができる．Bellman の著書[3]はこの adaptive 制御への案内書として書かれたものだが，これを通読しても，この非常に広い範囲の，興味をそそる，また重要な問題の解析的あるいは数値的な研究がまだ始まったばかりであるということがわかる．いままでに得られた主な結果は [3] [4] にある Aoki, Bellman-Kalaba のものの他にも Bellman-Kalaba [42] [43]，Freimer [44]，Kramer, Jr. [45]，Sakaguchi [46] などがある．

(c) 数理統計学への応用——問題の本質を極値演算をもった漸化関係式あるいは函数方程式に invariant imbedding するという DP の思想は，A. Wald の sequential analysis の仕事(1945 年)の根底にもあることで，一般に統計的逐次決定の理論に DP の技法が使えるだろうことは当然考えられるところである．α を助変数とする確率密度函数族 $\{\xi(\theta|\alpha)\}_\alpha$ が，密度函数 $\phi(x|\theta)$ をもつ観測値につき closed under sampling(略して c. u. s.)だというのは

$(\forall\alpha, x_1, \cdots, x_n)\,\exists\beta=\beta(\alpha, x_1, \cdots, x_n)$;

$$\xi(\theta|\beta)=\xi(\theta|\alpha)\prod_{i=1}^{n}\phi(x_i|\theta)\Big/\int\xi(\theta|\alpha)\prod_{i=1}^{n}\phi(x_i|\theta)\,d\theta$$

のことであると定義する [47]．(例：Beta 分布は 2 項分布観測値につき c. u. s. など)簡単のために 2-決定問題を考えよう．$W_i(\theta)$ $(i=1, 2)$ を θ が真の母数値のときに最終決定 d_i をなすことの loss とする．未知母数 θ に関する知識が $\xi(\theta|\alpha)$ のときに最終決定 d_1 または d_2 をなすか，または更に観測を続けるか何れかにきめなければならないから

$R(\beta)$……知識が $\xi(\theta|\beta)$ のときから出発して，最適政策による全 risk

とおくと明らかに

$$R(\beta)=\min\begin{bmatrix} d_1: \int W_1(\theta)\xi(\theta|\beta)d\theta \\ d_2: \int W_2(\theta)\xi(\theta|\beta)d\theta \\ C: c+\int R(\beta_x)dx\int \phi(x|\theta)\xi(\theta|\beta)d\theta \end{bmatrix}$$

という多段選択過程になる．ここに c は観測値 1 個の抽出費用，β_x は x を得た後の事後知識：

$$\xi(\theta|\beta_x)=\xi(\theta|\beta)\phi(x|\theta)\Big/\int\xi(\theta|\beta)\phi(x|\theta)d\theta$$

である．統計的決定理論をこのように眺めたゆき方は Blackwell-Girshik [48] や Weiss [49] の著書に強調されているし，興味ある論文として例えば [29] [47] [50] など多数がある．1952 年に始めて Robbins [51] によりその重要さが指摘された sequential design の問題は Robbins ら，Karlin ら [52]，Blackwell ら [53]，Chernoff [54]，および Feldman [55] により，主として stochastic approximation (例えば [56]) や two-armed-bandit problem の形で興味ある展開をみせてきている．

参　考　文　献

筆者が本誌「文献抄録」欄に抄録したものは，各論文末尾に { } として付記した．

[1] Bellman, R., "On the theory of dynamic programming," *Proc. Nat. Acad. Sci. U. S. A.*, Vol. 38 (1952), 716–719.

[2] ——, *Dynamic Programming*, Princeton University Press, 1957.

[3] ——, *Adaptive Control Processes: A Guided Tour*, Princeton University Press, 1961. {5巻2号, 127}.

[4] ——, *Applied Dynamic Programming*, Princeton University Press, 1962.

[5] ——, Glicksberg, I. & Gross, O., "Some aspects of the mathematical theory of control process," RAND Report 313, 1958. {4巻1号, 49}.

[6] ——, "The theory of dynamic programming" *Bull. Amer. Math. Soc.*, Vol 60 (1954), 503–516.

[7] ——, "Theory of dynamic programming," pp. 243–278 in [8].

[8] Beckenbach, E. F., (ed), *Modern Mathematics for the Engineer*, *Ser. I*, McGraw Hill Book Co., 1956.

[9] Dreyfus, S., "Dynamic Programming," Ackoff, R. L. (ed.), *Progress in Operations Research, Vol I*, John Wiley, 1961, 211–242.

[10] Bellman, R., "Functional equations and successive approximations in linear and non-linear programming," *Naval Res. Logist. Quart.*, Vol 7 (1960), 63–84. {4巻1号, 58}.

[11] Karush, W. and Vazsonyi, A., "Mathematical programming and employment scheduling," *Naval Res. Logist. Quart.*, Vol. 4 (1957), 297–320.

[12] Arrow, K. J., Harris, T. E. and Marschak, J., "Optimal inventory policy," *Econometrica*, Vol. 19 (1951), 250–272.

[13] Dvoretzky, A., Kiefer, J. and Wolfowitz, J., "On the optimal character of the (s, S) policy in inventory theory," *Econometica*, Vol. 21(1953), 586-596.

[14] Karlin, S., "Dynamic inventory policy with varying stochastic demands.," *Manag. Sci.*, Vol. 6(1960), 231-258. {4巻1号, 51}.

[15] Scarf, H., "Bayes solutions of the statistical inventory problem," *Ann. Math. Stat.*, Vol. 30(1959), 490-508. {3巻3号, 179}.

[16] ――, "Some remarks on Bayes solutions to the inventory problem," *Naval Res. Logist. Quart.*, Vol. 7(1960), 591-596. {5巻2号, 125}.

[17] Beckmann, M. J., "An inventory policy for repair parts," 同上誌, Vol. 6(1959), 209-220.{4巻4号, 248}.

[18] Arrow K. J., Karlin, S., and Scarf, H., *Studies in the Mathematical Theory of Inventory and Production*, Stanford Univ. Press, 1958.

[19] Karlin, S., "The mathematical theory of inventory processes," pp. 228-260 in [20].

[20] Beckenbach, E. F. and Hestenes, M. R., *(ed)*, *Modern Mathematics for the Engineer*, *Ser.* II, McGraw Hill Book Co., 1961.

[21] Gross, O. and Johnson, S. M., "Sequential minimax search for a zero of a convex function," *Math. Tables and Other Aids to Comp.*, Vol. 13(1959), 44-51. {4巻1号, 49}.

[22] Bellman, R. and Gluss, B., "On various versions of the defective coin problem," *Informaion and Control*, Vol. 4(1961), 118-131. {5巻4号, 241}.

[23] Gluss, B., "An optimal policy for determining a fault in a complex system," *Oper. Res.*, Vol. 7(1959), {4巻3号, 199}.

[24] Isbell, J. R., "an optimal search pattern," *Naval Res. Logist. Quart.*, Vol. 4(1957), 357-360. {4巻3号, 198}.

[25] Blachman, R. J., "Prolegomena to optimum discrete search procedure," 同上誌 Vol. 6(1959) 273-282. {4巻4号, 246}.

[26] Gluss, B., "Approximately optimal one dimensional search policies in which search costs vary through time," 同上誌 Vol 8(1961), 277-284. {5巻4号, 244}.

[27] Bellman, R., "Formulation of recurrence equations for shuttle process and assembly line," 同上誌 Vol. 4(1957), 321-334.

[28] Macqueen, J. and Miller, R. G. Jr., "Optimal persistence policies," *Oper. Res.*, Vol 8(1960), 362-380. {4巻2号, 142}.

[29] Lindley, D. V., "Dynamic programming and decision theory," *Appl. Stat.*, Vol. 10(1961), 39-51. {5巻4号, 242}.

[30] Sakaguchi, M., "Dynamic programming of some sequential sampling design," *J. Math. Analysis and Appl.*, Vol. 2(1961) 446-466.

[31] Howard, R. A., *Dynamic Programming and Markov Processes*, Technology Press and Wiley, New York, 1960. {5巻2号, 130}.

[32] Blackwell, D., "Discrete dynamic programming," *Ann. Math. Stat.*, Vol. 33(1962), 719-726.

[33] Herniter, J. D. and Magee, J. F., "Customer behavior as a Markov process," *Oper. Res.*, Vol. 9(1961), 105-122.

[34] Dreyfus, S. E., "Dynamic programming and the calculus of variations," *proc. 2nd Intern. conf. on Oper. Res.*, 1960, 142-149.

[35] Bellman, R. and Lehman, R. S., "A mathematical model of two independent processes," *RAND Memor.* (1954).

[36] Bellman, R., Glicksberg, I. and Gross, O., "On the bang-bang control problem," *Quart. Appl. Math.*, Vol. 14(1956), 11-18, {4巻1号, 48}

[37] Isaacs, R., "Differential games I, II, III, IV" *RAND Memor.*, no. 1398, 1399, 1411, 1486 respectively.

[38] Dresher, Tucker and Wolfe *(ed)*, *Contributions to the Theory of Games*, III, Princeton Univ. Press, 1957 の中の論文 413-435, 407-412, 393-406.

[39] Zachrisson, L. E., "A tank duel with game-theoretic implications," *Naval Res. Logist. Quart.*, Vol. 4(1957), 131-139. {3巻3号, 177}.

[40] Weiss, H. K., "Some differential games of tactical interest and the value of a supporting weapon system," *Oper. Res.*, Vol. 7(1959), 180-196. {3巻3号, 171}.

[41] Bellman, R. and Dreyfus, S., "Functional approximation and dynamic programming," *Math. Tables and Other Aids to Comp.*, Vol. 13(1959), 247-251.

[42] Bellman, R. and Kalaba, R., "On adaptive control processes— A mathematical theory of adaptive control processes," *Proc. Nat. Acad. Sci. U. S. A.*, Vol. 45(1959), 1288-1290. {3巻, 3号, 180}.

[43] ———, "On adaptive control processes" *IRE Nat. Conv. Record*, Part 4, 1959, 3-11.

[44] Freimer, M., "A dynamic programming approach to adaptive control processes," 同上誌, Part 4, 1959, 12-17.

[45] Kramer Jr., J. D. R., "On control of linear systems with time lags," *Information and Control*, Vol. 3(1960), 299-236.

[46] Sakaguchi, M., "Information pattern, learning structure and optimal decision rule," to appear in *Information and Control*.

[47] Wetherill, G. B., "Bayesian sequential analysis," *Biometrika*, Vol. 48 (1961), 281-292. {5巻3号, 199}.

[48] Blackwell, D. and Girshick, M. A., *Theory of Games and Statistical Decisions*, Wiley, New York, 1954.

[49] Weiss, L., *Statistical Decision Theory*, Wiley, New York, 1961.

[50] Lindley, D, V., "Binomial sampling schemes and the concept of information," *Biometrika*, Vol. 44(1957), 179-186.

[51] Robbins, H., "Some aspects of the sequential design of experiments," *Bull. Amer. Math. Soc.*, Vol. 58(1952), 527-535.

[52] Bradt, R. N., Johnson, S. M. and Karlin, S,. "On sequential designs for maximizing the

sum of *n* observations," *Ann. Math. Stat.*, Vol. 27(1956),1060-1074.

[53] Blackwell, D. and Hodges Jr., L., "Design for the control of selection bias," *Ann. Math. Stat.*, Vol. 29(1958), 449-460.

[54] Chernoff, H., "Sequential design of experiments," *Ann. Math. Stat.*, Vol. 30(1959), 755-770. {4巻1号, 47}.

[55] Feldman, D., "Contributions to the two-armed-bandit problem," *Ann. Math. Stat.*, Vol. 33 (1962), 847-856.

[56] 北川敏男,「推測過程論」岩波応用数学講座のうち, 岩波書店, 1958.

ニ ュ ー ス Goodeve 卿来日. 国際 OR 学会連合初代の会長であった Sir Charles F. Goodeve (The British Iron and Steel Research Association) が, 日本鉄鋼協会の招きで来年3月に来日される.

★ **IFORS だより**★ Oslo 大会への代表選出についての Morse 教授の手紙は別項のとおり. その後の同教授よりの来信によると, プログラム委員会へ申込んだ論文数は 58 [米 21(5), 仏 15(2), 英 9(4), ドイツ及びスイス 3(1), オランダ 3(1), ノルウェー(2), スウェーデン 2, カナダ 1, 日 1, 伊 1, デンマーク 1(1), ポルトガル 1, イスラエル(1) (カッコ内は各 session や study group の Chairman の数で概数)] となっている. なおアメリカ OR 学会では春に Cleveland で大学院学生の発表会を開き, 最優秀とみとめられた者は Oslo 大会の代表とするとともに, その費用を"学割"にするという賞をつけるそうである.