Chapter 8

時間を通じての最大化問題 III: ダイナミック・プログラミング

最適制御問題、変分法とならんで、時間を通じての最適化に用いられるのが、動的計画法である。最終的には、どの手法でも最適解が差分方程式の形で表現される。しかし、解に至る考え方は、かなり異なる。最適制御問題では、すべての変数をわけへだてなく扱うことにより静学的最適値問題に帰着させ、最大値原理のような一般原理に持ち込む。変分法では、差分（あるいは微分）を含む関数方程式の最大化問題からオイラー方程式のような陰関数的差分（あるいは微分）方程式を導出する。これに対して、動的計画法では、時間を通じての最適値問題であるという特性を意識して、解法をさぐる。実際には、「ある期間を通じての最適問題の解として最適政策は、元々の問題を部分期間に区切った部分最適問題の解を一部として持つ」という最適性の原理から、ベルマン方程式という関数方程式を得る。

ダイナミック・プロググロミングは、考え方がわかりやすく、不確実性を含む問題についても適用しやすく、経済学での応用例も多い。とはいえ、関数方程式としてのベルマン方程式が、最大値原理やオイラー方程式にくらべて解きにくいという欠点がある。要は使い方の問題である。

8.1 ダイナミック・プログラミングの問題

ダイナミック・プログラミングには、典型的な問題設定というものはない。ここでは、最適制御と同じような問題を考えておこう。つまり、

$$\text{maximize} \quad \sum_{t=0}^{T-1} U(x(t), u(t)) + \nu(x(T)) \tag{8.1}$$

subject to

$$x(t+1) - x(t) = f(x(t), u(t)), \quad (t = 0, \ldots, T-1) \tag{8.2}$$

$$g(x(t), u(t)) \geq 0, \quad (t = 0, \ldots, T-1) \tag{8.3}$$

$$x(0) = x_0 \tag{8.4}$$

8.2 後ろ向き帰納法

さて、後ろ向き帰納法という考え方で、ベルマンの最適性の原理を生かして上の問題を解いてみよう。ただし、解くといっても、最適の必要条件が具体的に求まるわけではない。

後ろ向き帰納法とは、「最適だったなら、前の期の制御はこうでなくてはならない。そのとき、前の期の状態はこうである。もしそうなら、さらに前の期の制御はこうでなくてはならない。そのとき状態はこうだ......」というような議論は初期時点までつづけ、最適解を特徴づけるというやりかたである。

今、最適制御$u(0)^*, u(1)^*, \ldots, u(T-1)^*$が定まっているとする。

T期　T期では、$T-1$の制御$u(T-1)^*$により状態方程式から状態変数$x(T)$が定まっており、制御を特徴づける条件はない。

$T-1$期　前期のu^*_{T-2}により状態方程式から状態変数$x(T-1)$が決まっている。しかし、$u(T-1)$は、来期（最終期T期）の状態を定め、最終期関数$\nu(x(T))$を動かすから、$u(T-1)^*$は次の条件付き最大化問題の解でなくてはならない。

$$\text{maximize} \quad U(x(T-1), u(T-1)) + \nu(x(T)) \tag{8.5}$$

$$\text{subject to}$$

$$g(x(T-1),u(T-1)) \geq 0 \tag{8.6}$$

$$x(T) - x(T-1) = f(x(T-1),u(T-1)) \tag{8.7}$$

これは、最適制御$u(0)^*, u(1)^*, \ldots, u(T-1)^*$は$T-1$期から$T$期にかけての部分問題の解を持たなくてはならないという最適性の原理そのものである。上記の特徴づけは、$u(T-1)^*$は$T-2$期以前に最適制御をとった結果に達した状態$x(T-1)^*$の関数になっていることを意味する。これを$u(T-1)^* = u_{T-1}(x(T-1)^*)$と書く。なお、

$$V(x(T-1)) = U(x(T-1),u(T-1)) + \nu(x(T-1) + f(x(T-1),u_{T-1}(x(T-1)))) \tag{8.8}$$

と書かれる関数は$T-1$期に$x(T-1)$であったときに、$T-1$期の制御を動かしてもたらされる最大の目標値である。

$T-2$期以前、1期以降　これらも同様に、t期の最適制御が前期までのの最適制御の結果もたらされる今期状態$x(t)^*$の関数として表現されるが、その場合つぎのように再帰的に定義される状態評価関数

$$V(x(t)) = \max_{u(t)} \{U(x(t),u(t)) + V(x(t+1)) \mid g(x(t),u(t)) \geq 0,\ x(t+1) - x(t) = f(x(t),u(t))\} \tag{8.9}$$

$$V(x(T)) = \nu(x(T)) \tag{8.10}$$

を用いて、

$$V(x(t)) = U(x(t),u(t)) + V(x(t+1), x(t) + f(x(t),u(t))) \tag{8.11}$$

という方程式の解であると考えるのが、最適性の原理である。

0期　ここでは初期最適制御$u(0)$が所与の$x(0) = x_0$に依存して決まる。ただし、そこでの決まり方は、状態評価関数を用いた方程式によることは同じである。

具体的には、0期を出発点にして、逐次的に最適経路がもとまっていく。

8.3 基本的再帰関係式

前の節では、最終時点から状態評価関数を逆算する形で、最適解を特徴づけようとした。ここでは、最適制御問題の部分問題の目標関数の最適値として状態評価関数を定義し、これを用いて最適解を特徴づけよう。

$V(x(t))$を、次の問題

$$\text{maximize} \quad \sum_{\tau=t}^{T-1} U(x(\tau),u(\tau)) + \nu(x(T)) \tag{8.12}$$

$$\text{subject to}$$

$$x(\tau+1) - x(\tau) = f(x(\tau),u(\tau)), \quad (\tau = t, \ldots, T-1) \tag{8.13}$$

$$g(x(\tau),u(\tau)) \geq 0, \quad (\tau = t, \ldots, T-1) \tag{8.14}$$

$$x(t) = x_t \tag{8.15}$$

の最適値とする。ベルマンの最適性の原理とは、次の関数方程式が満たされることを要求するものと考えられる。

$$V(x(t)) = \max_{u(t)} \{U(x(t), u(t)) + V(x(t+1) + f(x(t), u(t))) \mid g(x(t), u(t)) \geq 0\} \tag{8.16}$$

これは、境界条件

$$V(x(t)) = \nu(x(t)), \quad x(0) = x_0 \tag{8.17}$$

によって、$\{V(x(t))\}_{t=0,\ldots,T}$が決まる。各時点の最適制御$u(t)$は、$V(x(t))$を用いて、

$$\begin{aligned} &\text{maximize} && U(x(t), u(t)) + V(x(t) + f(x(t), u(t)) \\ &\text{subject to} && g(x(t), u(t)) \geq 0 \end{aligned} \tag{8.18}$$

の解としてもとまる。最適な状態は、さらに状態方程式により定まる。

8.4 応用例：最適成長

生産関数を

$$Y_t = F(K_t) \tag{8.19}$$

とする。これは、資本用役を投入とする生産関数である。生産物は消費か、資本形成（投資）にまわされる。つまり、

$$Y_t = C_t + I_t \tag{8.20}$$

また、資本の増加は次の式による。

$$K_{t+1} - K_t = I_t - \delta K_t \tag{8.21}$$

$$K_0 = \text{given} \tag{8.22}$$

δは資本減耗率とする。計画期間は0期からT期までとし、最終のT期の処分から得られる効用単位の価値を$\nu(K_T)$とする。各期の効用は消費により$u(C_t)$で決まる。効用の割引因子をβとして、計画主体の目的は

$$\sum_{t=0}^{T-1} \beta^t u(C_t) + \beta^T \nu(K_T) \tag{8.23}$$

を生産関数と処分式、資本蓄積方程式の制約の下で最大化する。

生産関数は

$$F(K_t) = K_t^a, \quad 0 < a < 1 \tag{8.24}$$

$$u(C_t) = \log C_t \tag{8.25}$$

$$\nu(K_T) = A_T + B_T \log K_T, \quad 0 < B_T \tag{8.26}$$

また、簡単化のために$\delta = 1$と仮定する。

以上の問題をダイナミック・プログラミングによって解いてみよう。まず、T期では、選択の余地はない。状態評価関数は、境界条件によって

$$V(K_T) = \beta^T(A_T + B_T \log K_T) \tag{8.27}$$

次に、$T-1$期では

$$V(K_{T-1}) = \max_{C_{T-1}} \log C_{T-1} + \beta^T(A_T + B_T \log(K_{T-1}^a - C_{T-1})) \tag{8.28}$$

この段階で、最適な制御としての$T-1$期の消費量が

$$C^*_{T-1} = \frac{1}{1+\beta B_T} K^a_{T-1} \tag{8.29}$$

これから、

$$V(K_{T-1}) = \beta^{T-1}(A_{T-1} + B_{T-1}\log K_{T-1}) \tag{8.30}$$

ここで、

$$A_{T-1} = \log\frac{1}{1+\beta B_T} + \beta A_T + \beta B_T \log\frac{\beta B_T}{1+\beta B_T} \tag{8.31}$$

$$B_{T-1} = a(1+\beta B_T) \tag{8.32}$$

となる。$T-2, T-3, \ldots, 0$ についても同様の関係式が得られ、一般的に

$$C^*_t = \frac{1}{1+\beta B_{t+1}} K^a_t \tag{8.33}$$

$$A_t = \log\frac{1}{1+\beta B_{t+1}} + \beta A_{t+1} + \beta B_{t+1}\log\frac{\beta B_{t+1}}{1+\beta B_{t+1}} \tag{8.34}$$

$$B_t = a(1+\beta B_{t+1}) \tag{8.35}$$

上の式においてA_T, B_T所与を考慮すると、逐次代入によって、C^*_tが求まることがわかった。