スケジュールの最適化理論

@ 第 4 回理論生物学 夏の学校

大槻 久

September 10th, 2024

総研大・統合進化科学研究センター

概要

- 動的最適化とは
- 動的最適化の基礎理論
  — 最適性原理と Bellman（ベルマン）方程式
  — 例：コロニーの最適繁殖戦略
- 連続モデルの動的最適化
  — 例：最適植林スケジュール
  — Hamilton-Jacobi-Bellman 方程式
  — Pontryagin（ポントリャーギン）の最大原理

動的最適化とは

動的最適化とは

最適化 (optimization)

意思決定できる状況で、ある目的関数 J を最大（最小）化すること。

例　$J = -u^2 + 2u$ を最大化する u を求めよ。

動的最適化 (dynamic optimization)

各時刻 t ごとに意思決定できる状況で、ある目的関数 J を最大（最小）化すること。

例　$\dot{x}(t) = u(t)$ の時、$J = \int_0^1 \left(x(t) - \frac{u(t)^2}{2} \right) \mathrm{d}t$ を最大化する $u(t)$ は？

動的最適化のほうがずっと難しい。

動的最適化と数理生物学

動的最適化は生物学において様々な応用がある、たとえば、

- 生物がいつ成長から繁殖へ切り替えるか。
- 生物資源の成長に合わせていかに持続可能な形で利用するか。
- 絶滅危惧種をどのタイミングでどれだけ保全するか。
- 感染症をどのステージでどれだけ抑え込むか。
- 疾病の進行段階に応じてどのように治療するか。
- モルフォゲンを発生のどの段階でどれだけ産生するか。

動的最適化の書籍

主に制御工学の書籍が多く、数理生物学者が学ぶ機会は限られている。

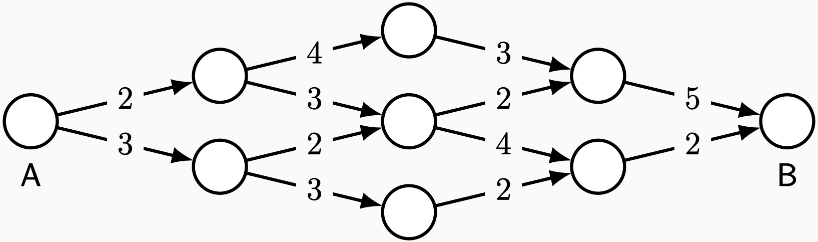
数理生物学関連の書籍

- 『数理生物学入門』巌佐 庸（12 章）
- 『進化生態学入門』山内 淳（3 章）
- 『「行動・進化」の数理生物学』日本数理生物学会（4 章）

動的最適化の基礎理論

例題

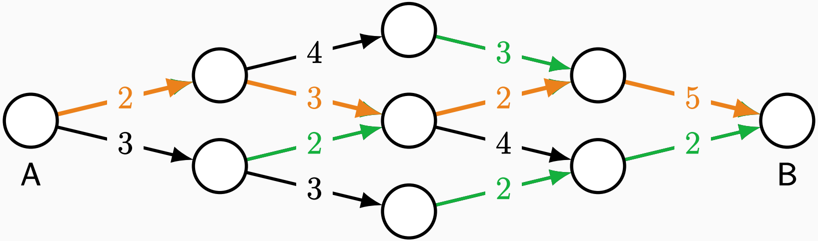
各道路に通行料が定められている時、地点 A から B に最安で行く方法は？



例題

貪欲法 (greedy method) を試す。

- 各 node で最安経路を用いる。

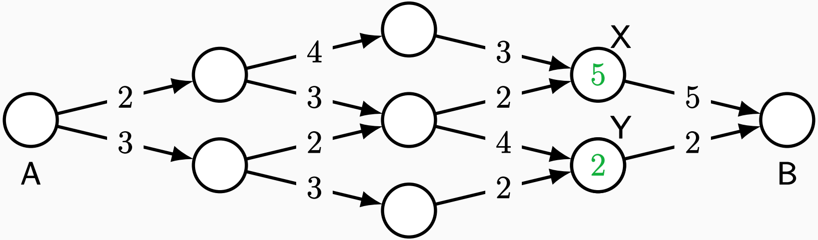


しかし 2-3-2-5($= 12$) は最適経路ではない（最適は一番下の 3-3-2-2($= 10$)）。

例題

貪欲法の問題点 ⇒ 目先の利益に飛びつき将来の損失を考慮できていない。

そこで各 node の将来性を考えてみる。

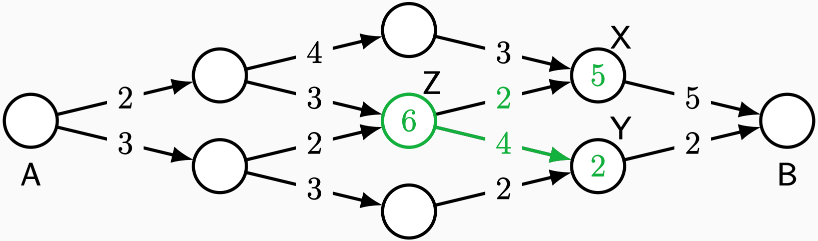


- node X に着くと、次に 5 払うことが確定 ⇒ X の「将来コスト」は 5
- node Y に着くと、次に 2 払うことが確定 ⇒ Y の「将来コスト」は 2

例題

そこで「即時」と「将来」のコストの和を最小化する行動を考える。

- node Z での意思決定
  - 上に行くと、コストは（即時 2）と（将来 5）の和。
  - 下に行くと、コストは（即時 4）と（将来 2）の和。⇒ こちらが Z での最適行動。
- Z 以降の最小コストは $4 + 2 = 6$ ⇒ Z の「将来コスト」を 6 と定義。

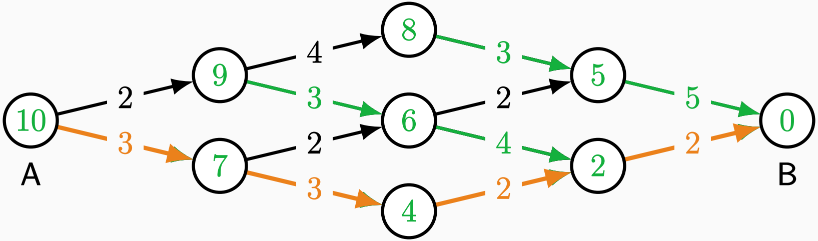


例題

この手順を繰り返すと、各 node の将来コストが計算できる。

さらに各 node で「即時+将来」コストを最小にする最適行動が定義できる。

各 node での最適行動をつなげば、大域的な最適化問題の解が手に入る。



このような手法を動的計画法 (dynamic programming) と呼ぶ。

数学的定式化

価値関数 (value function)

各 node s に対し、s 以降で最適行動を取り続けると得られる最大の利益を $V(s)$ で表し、s の価値関数と呼ぶ。

Remarks:

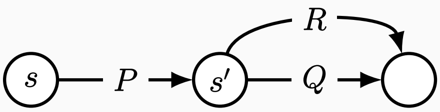
- コスト最小化問題 ⇒ コストに負符号をつければ最大化問題となる。
- 一般に最大値 $\max$ が存在するとは限らない ⇒ その時は上限値 $\sup$ で定義。

数学的定式化

最適性原理 (principle of optimality)

s から始まる最適化問題の最適経路において、その途中にある s' 以降の経路は、s' から始まる最適化問題の最適経路となっている。

証明：
背理法。s から始まる問題の最適経路を $P+Q$ と分割する（右図）。Q が s' から始まる問題の最適経路でないとすると、より優れた経路 $R(\neq Q)$ が取れる。すると s から始まる問題において経路 $P+R$ は経路 $P+Q$ よりも優れており矛盾。■



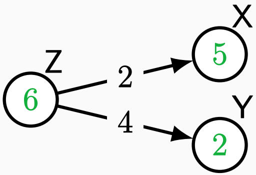
Bellman 方程式

Bellman 方程式 (Bellman equation)

node s で選択 a をした結果得られる、即時利益を r_{sa}、遷移後の状態を s' とする（注：s' は a に依存する）。この時、価値関数 V は

$$V(s) = \max_a \{r_{sa} + V(s')\}$$

を満たす。これを Bellman 方程式と呼ぶ。



$$\underbrace{V(\mathrm{Z})}_{6} = \min\{2 + \underbrace{V(\mathrm{X})}_{5}, 4 + \underbrace{V(\mathrm{Y})}_{2}\}$$

Remark: 最大値がなければ上限 sup を使う。以下では最大値 max の存在を仮定。

生物学の例

コロニーの最適繁殖戦略

初期状態でワーカー w_0 個体からなる社会性昆虫コロニーを考える。
各 $t\,(=1,\cdots,T)$ 期では「繁殖 (R)」「成長 (G)」のいずれかを選ぶことができ、

- 繁殖を選ぶとワーカー数はそのまま（$w_t = w_{t-1}$）だが、
  新女王を $q_t = rw_{t-1}$ 匹生産できる。
- 成長を選ぶとワーカー数は増加（$w_t = gw_{t-1}$）するが、
  新女王は生産できない（$q_t = 0$）。

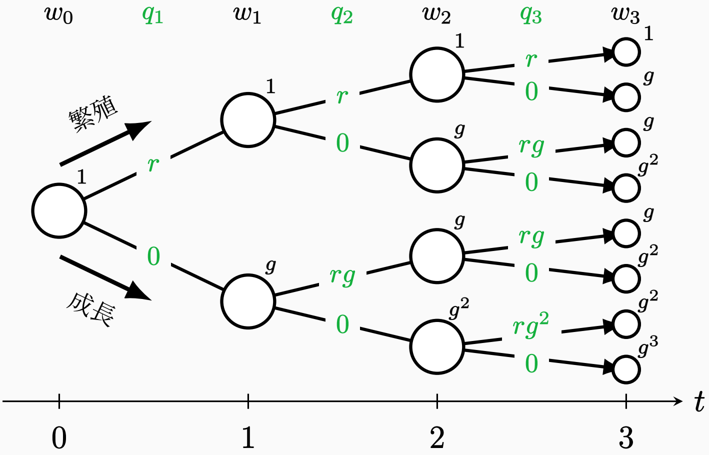
目的関数を

$$J = \text{（生産された新女王の総数）} = q_1 + \cdots + q_T \quad \longrightarrow \max$$

とするとき、最適繁殖スケジュールを求めよ。

生物学の例

$w_0 = 1$ とした場合：



Bellman 方程式

系の状態 s は、$s = (\text{時刻}, \text{ワーカー数}) = (t, w)$ で記述される。

Bellman 方程式

$$V(t, w) = \max\{\underbrace{rw + V(t+1, w)}_{\text{繁殖}}, \underbrace{0 + V(t+1, gw)}_{\text{成長}}\} \qquad (t = 0, \cdots, T-1)$$

$$V(T, w) = 0$$

Bellman 方程式の解法

後ろの時間から解く。

時刻 T における意思決定

$$V(T-1,w) = \max\{rw + \underbrace{V(T,w)}_{=0},\ 0 + \underbrace{V(T,gw)}_{=0}\}$$
$$= \max\{\underbrace{rw}_{\text{繁殖}},\ 0\} = rw$$

よって繁殖が最適。

Bellman 方程式の解法

時刻 $T-1$ における意思決定

$$V(T-2,w) = \max\{rw + \underbrace{V(T-1,w)}_{=r\cdot w},\ 0 + \underbrace{V(T-1,gw)}_{=r\cdot gw}\}$$

$$= \max\{2rw,\ grw\}$$

よって、

- $g > 2$ の時、成長が最適で、$V(T-2,w) = \max\{2rw,\ \underbrace{grw}_{\text{成長}}\} = grw$
- $g < 2$ の時、繁殖が最適で、$V(T-2,w) = \max\{\underbrace{2rw}_{\text{繁殖}},\ grw\} = 2rw$

Bellman 方程式の解法

時刻 $T-2$ における意思決定（$g>2$ の場合）

$$V(T-3,w)=\max\{rw+\underbrace{V(T-2,w)}_{=gr\cdot w},\ 0+\underbrace{V(T-2,gw)}_{=gr\cdot gw}\}$$

$$=\max\{(1+g)rw,\ \underbrace{g^2rw}_{\text{成長}}\}=g^2rw$$

よって成長が最適。

Bellman 方程式の解法

時刻 $T-2$ における意思決定（$g<2$ の場合）

$$
\begin{aligned}
V(T-3, w) &= \max\{rw + \underbrace{V(T-2, w)}_{=2r\cdot w},\ 0 + \underbrace{V(T-2, gw)}_{=2r\cdot gw}\} \\
&= \max\{3rw,\ 2grw\}
\end{aligned}
$$

よって、

- $g > 3/2$ の時、成長が最適で、$V(T-3, w) = \max\{3rw, \underbrace{2grw}_{\text{成長}}\} = 2grw$
- $g < 3/2$ の時、繁殖が最適で、$V(T-3, w) = \max\{\underbrace{3rw}_{\text{繁殖}}, 2grw\} = 3rw$

Bellman 方程式の解

最適解は以下のような形になる。

	T	$T-1$	$T-2$	$T-3$	$\cdots$
$g > 2$	繁	成	成	成	$\cdots$
$3/2 < g < 2$	繁	繁	成	成	$\cdots$
$4/3 < g < 3/2$	繁	繁	繁	成	$\cdots$
$5/4 < g < 4/3$	繁	繁	繁	繁	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

演習：上を示せ。

生物学的示唆

しばらく成長に投資してから、最後のほうで繁殖するスケジュールが最適である。

たとえば $g > 2$ の時、価値関数は

$$V(1, w) = g^{T-2} r w$$

であるから、$t = 1$ での意思決定では

$$V(0, w_0) = \max\{\underbrace{r w_0 + g^{T-2} r w_0}_{\text{繁殖}}, \underbrace{0 + g^{T-1} r w_0}_{\text{成長}}\}$$

なる選択を迫られている。

- 繁殖すれば即時に新女王 $r w_0$ を得られるが、成長すれば新女王は 0。
- 繁殖すれば将来価値は $g^{T-2} r w_0$ だが、成長すればその g 倍（莫大な利益！）。

機械学習との関連

Bellman 方程式は機械学習（強化学習）でよく用いられる。

$$V(s) = \max_a \{r_{sa} + \gamma V(s')\} \qquad (\gamma：割引率)$$

$V(s)$ は状態 s の状態価値関数と呼ばれる。

$Q(s,a) = r_{sa} + \gamma V(s')$ は、行動価値関数と呼ばれ、状態 s で行動 a をとることの価値である。

学習機械にとって r_{sa} は未知なので、r_{sa} を受け取る度に $Q(s,a)$ を

$$Q(s,a) \longleftarrow (1-\alpha)Q(s,a) + \alpha[r_{sa} + \gamma V(s')] \qquad (\alpha：学習率)$$

で更新する。$V(s')$ も未知なので $\max_{a'} Q(s',a')$ で代用する：

$$Q(s,a) \longleftarrow (1-\alpha)Q(s,a) + \alpha[r_{sa} + \gamma \max_{a'} Q(s',a')]$$

これを Q 学習 (Q-learning) と呼ぶ。

連続モデルの動的最適化

生物学の例

最適植林スケジュール

森林を管理し生態系サービスを享受したいとする。

時刻 t における植林量を $u(t)$ (≥ 0) とした時、森林面積 $x(t)$ は

$$\dot{x}(t) = u(t) - dx(t) \qquad (d：死亡率、初期値\ x(0) = x_0)$$

に従い変化する。この時、目的関数

$$J = \int_0^T \Bigg(\underbrace{x(t)}_{\substack{\text{生態系}\\\text{サービス}}} - \underbrace{c\frac{u(t)^2}{2}}_{\text{植林コスト}} \Bigg) \mathrm{d}t \quad \longrightarrow \max$$

を最大にする植林スケジュール $u : [0, T] \to \mathbb{R}_{\geq 0}$ を求めよ。

数学的定式化

時間 ： $t \in [0, T]$ (T は固定)
状態変数 (state variable) ： $x(t) \in X$
制御変数 (control variable) ： $u(t) \in U$
状態変数の従う微分方程式 ： $\dot{x}(t) = f(t, x(t), u(t))$ （初期値 $x(0) = x_0$）
目的関数 (objective function) ：

$$J = \int_0^T L(t, x(t), u(t))\mathrm{d}t + K(x(T)) \quad \longrightarrow \max$$

Remark: 上の J を Bolza 型目的関数という。$K = 0$ なら Lagrange 型、$L = 0$ なら Mayer 型と呼ばれる。

Remark: x, u はベクトル値 $\boldsymbol{x}, \boldsymbol{u}$ でも良い。

Dynamic Programming による定式化

価値関数を以下のように定める：

$$V(t,x) = \max_{\substack{u:[t,T]\to U \\ \text{under } x(t)=x}} \left\{ \int_t^T L(s,x(s),u(s))\mathrm{d}s + K(x(T)) \right\}$$

『$V(t,x)$ とは、時刻 t から初期値 x で始めた時に得られる、以降の最大利益』

Remark：以下、$\max$ の存在を仮定。

問：価値関数の従う式は？

Dynamic Programming による定式化

以下、V は C^1 級であることを仮定。微小時間 Δt を用いて

$$
\begin{aligned}
V(t,x) &= \max_{\substack{u:[t,T]\to U \\ \text{under } x(t)=x}} \left\{ \left(\int_t^{t+\Delta t} + \int_{t+\Delta t}^{T} \right) L(s,x(s),u(s))\mathrm{d}s + K(x(T)) \right\} \\
&\simeq \max_{\substack{u:[t,T]\to U \\ \text{under } x(t)=x}} \left\{ L(t,x,u)\Delta t + \int_{t+\Delta t}^{T} L(s,x(s),u(s))\mathrm{d}s + K(x(T)) \right\} \\
&= \max_{u\in U} \left[L(t,x,u)\Delta t + \max_{\substack{u:[t+\Delta t,T]\to U \\ \text{under} \\ x(t+\Delta t)=x+f(t,x,u)\Delta t}} \left\{ \int_{t+\Delta t}^{T} L(s,x(s),u(s))\mathrm{d}s + K(x(T)) \right\} \right] \\
&= \max_{u\in U} \left[L(t,x,u)\Delta t + V(t+\Delta t, x+f(t,x,u)\Delta t) \right]
\end{aligned}
$$

Dynamic Programming による定式化

$$V(t,x) \simeq \max_{u\in U}\left[L(t,x,u)\Delta t + V(t+\Delta t, x + f(t,x,u)\Delta t)\right]$$

Taylor 展開を行う（以下、下添字はその変数による偏微分を表す）。

$$V(t,x) \simeq \max_{u\in U}\left[L(t,x,u)\Delta t + V(t,x) + V_t(t,x)\Delta t + V_x(t,x) f(t,x,u)\Delta t\right]$$

$$0 \simeq \max_{u\in U}\left[L(t,x,u)\Delta t + V_t(t,x)\Delta t + V_x(t,x) f(t,x,u)\Delta t\right]$$

$$0 = \max_{u\in U}\left[L(t,x,u) + V_t(t,x) + V_x(t,x) f(t,x,u)\right]$$

$$-V_t(t,x) = \max_{u\in U}\left[L(t,x,u) + V_x(t,x) f(t,x,u)\right]$$

Hamilton-Jacobi-Bellman 方程式

Hamilton-Jacobi-Bellman 方程式 (HJB equation)

価値関数 V の満たす偏微分方程式

$$-V_t(t,x) = \max_{u \in U} \left[L(t,x,u) + V_x(t,x) f(t,x,u) \right]$$

を Hamilton-Jacobi-Bellman 方程式 (HJB equation) と呼ぶ。

境界条件：定義より $V(T,x) = K(x)$ である。

Remark: 方程式の名前は解析力学の Hamilton-Jacobi 方程式に因む。

HJB 方程式の解と最適制御の関係

HJB 方程式を解くことができれば、最適解 u^* を見つけられる。

定理

HJB 方程式の解 $V(t,x)$ が存在したと仮定する。
この時、$u^*(t)$ が最適制御であることの必要十分条件は、

$$\begin{cases} \dot{x}^*(t) = f(t, x^*(t), u^*(t)), \quad x^*(0) = x_0 & (x^* \text{ は } u^* \text{ の下で解}) \\ u^*(t) \in \operatorname*{argmax}_u \left[L(t, x^*(t), u) + V_x(t, x^*(t)) f(t, x^*(t), u) \right] & (u^* \text{ は常に最適}) \end{cases}$$

を満たすことである。

Remark：記号 argmax_u は最大値を与える u の集合を表す。

しかし HJB 方程式は偏微分方程式なので、解くことは一般には難しい。

例

先程の植林スケジュールモデル。

$$f(t,x,u) = u - dx,\ L(t,x,u) = x - c\frac{u^2}{2},\ K(x) = 0$$

HJB 方程式は

$$V_t(t,x) = -\max_{u\in U}\left[\left(x - c\frac{u^2}{2}\right) + V_x(t,x)(u - dx)\right]$$

境界条件は $V(T,x) = 0$

別の方法

HJB 方程式の解 $V(t,x)$ が存在したと仮定する。
この時、$u^*(t)$ が最適制御であることの必要十分条件は

$$\begin{cases} \dot{x}^*(t) = f(t, x^*(t), u^*(t)), \quad x^*(0) = x_0 & (x^* \text{ は } u^* \text{ の下で解}) \\ u^*(t) \in \operatorname*{argmax}_u \left[L(t, x^*(t), u) + V_x(t, x^*(t)) f(t, x^*(t), u) \right] & (u^* \text{ は最適}) \end{cases}$$

を満たすことである。

上の定理は、$V(t,x)$ 自体が分からなくても、$V_x(t, x^*(t))$ という t の関数さえ分かれば $u^*(t)$ が分かることを示唆している。

したがって

$$\lambda(t) := V_x(t, x^*(t))$$

とおいて、これがどのような関数かを調べることにする。

別の方法

$\lambda(t) = V_x(t, x^*(t))$ に関して分かること（その１）。

境界条件より $V(T, x) = K(x)$ であった。

よって

$$\lambda(T) = V_x(T, x^*(T)) = K_x(x^*(T))$$

である。

別の方法

$\lambda(t) = V_x(t, x^*(t))$ に関して分かること（その２）。

Remark: 以下、V が C^2 級であることを仮定する。

$\lambda(t)$ の時間発展は

$$
\begin{aligned}
\dot{\lambda} &= \frac{\mathrm{d}}{\mathrm{d}t} V_x(t, x^*(t)) \\
&= V_{xt} + V_{xx}\,\dot{x}^*(t) \\
&= V_{xt} + V_{xx} f
\end{aligned}
$$

次で証明する補題より、$V_{xt} + V_{xx} f = -(L_x + \lambda f_x)$ が分かるので、

$$
\dot{\lambda} = -(L_x + \lambda f_x)
$$

別の方法

補題：$V_{xt} + V_{xx} f = -(L_x + \lambda f_x)$ である。

証明：HJB 方程式の導出で

$$0 = \max_{u \in U} \left[L(t, x, u) + V_t(t, x) + V_x(t, x) f(t, x, u) \right]$$

であった。右辺の四角カッコ内を $g(t, x, u)$ おくと、u^* の x^* に対する最適性から $0 = g(t, x^*(t), u^*(t))$ であり、一般の x に対しては $u^*(t)$ は最適とは限らないから $0 \geq g(t, x, u^*(t))$ である。

したがって、x の関数として見た時 $g(t, x, u^*(t))$ は $x = x^*(t)$ で最大値を取るから、特にその x-偏微分の値は 0 であり、

$$\left. \frac{\partial}{\partial x} g(t, x, u^*(t)) \right|_{x = x^*(t)} = L_x + V_{xt} + V_{xx} f + \underbrace{V_x}_{= \lambda} f = 0 \qquad \blacksquare$$

まとめ

HJB 方程式の C^2 級の解 $V(t,x)$ が存在したと仮定する。
この時、$u(t)$ が最適解である必要条件は、

$$
\begin{cases}
\dot{x}^*(t) = f(t, x^*(t), u^*(t)) & (x^* \text{ の発展式}) \\
x^*(0) = x_0 & (x^* \text{ の初期条件}) \\
\dot{\lambda}(t) = -\left[L_x(t, x^*(t), u^*(t)) + \lambda(t) f_x(t, x^*(t), u^*(t))\right] & (\lambda \text{ の発展式}) \\
\lambda(T) = K_x(x^*(T)) & (\lambda \text{ の終端条件}) \\
u^*(t) \in \operatorname*{argmax}_u \left[L(t, x^*(t), u) + \lambda(t) f(t, x^*(t), u)\right] & (u^* \text{ は常に最適})
\end{cases}
$$

なる $x^*(t)$ と $\lambda(t)$ が存在することである。

Hamiltonian と呼ばれる関数を $H(t,x,u,\lambda) := L(t,x,u) + \lambda(t) f(t,x,u)$ と定義すると、上の条件はよりコンパクトに次のように書ける。

まとめ

HJB 方程式の C^2 級の解 $V(t,x)$ が存在したと仮定する。

この時、$u(t)$ が最適解である必要条件は、

$$
\begin{cases}
\dot{x}(t) = H_\lambda(t, x(t), u(t), \lambda(t)), & x(0) = x_0 & (x \text{ の従う式}) \\
\dot{\lambda}(t) = -H_x(t, x(t), u(t), \lambda(t)), & \lambda(T) = K_x(x(T)) & (\lambda \text{ の従う式}) \\
u(t) \in \underset{u}{\operatorname{argmax}}\, H(t, x(t), u(t), \lambda(t)) & & (u \text{ は常に最適})
\end{cases}
$$

なる $x(t)$ と $\lambda(t)$ が存在することである。ただし H は Hamiltonian

$$H(t, x, u, \lambda) = L(t, x, u) + \lambda(t) f(t, x, u)$$

さらに、変分法を用いた全く別の証明方法によって、上の下線部を除けることが分かっている。

Pontryagin の最大原理

Pontryagin の最大原理 (Pontryagin's maximum principle)

$u(t)$ が最適解であることの必要条件は、

$$
\begin{cases}
\dot{x}(t) = H_\lambda(t, x(t), u(t), \lambda(t)), & x(0) = x_0 & (x \text{ の従う式}) \\
\dot{\lambda}(t) = -H_x(t, x(t), u(t), \lambda(t)), & \lambda(T) = K_x(x(T)) & (\lambda \text{ の従う式}) \\
u(t) \in \operatorname*{argmax}_u H(t, x(t), u(t), \lambda(t)) & & (u \text{ は常に最適})
\end{cases}
$$

なる $x(t)$ と $\lambda(t)$ が存在することである。ただし H は Hamiltonian

$$H(t, x, u, \lambda) = L(t, x, u) + \lambda(t) f(t, x, u)$$

λ は共役状態変数 (costate variable)、もしくは随伴変数 (adjoint variable)（もしくは補助変数）と呼ばれる。

Remarks

- H を Hamiltonian と呼ぶ理由は、解析力学における正準方程式
$$\begin{cases} \dot{q} = H_p \\ \dot{p} = -H_q \end{cases}$$
と同じ形をしているからである。
- 冷戦下にあって米国の Bellman が動的計画法を生み出したのと同時期かつ独立に、ソ連の Pontryagin は最大原理を生み出した。
  — 動的計画法は最適解の必要十分条件を与えるが、HJB 方程式を解くことは一般には難しい。
  — 最大原理は最適解の必要条件を与えるが、解くことは（HJB 方程式に比べると）相対的には容易である。

共役状態変数の直観的解釈

もともとは $\lambda(t) = V_x(t, x^*(t))$

$V(t,x)$ は状態 (t,x) の将来価値を表すのだった。よってその x-偏微分である $V_x(t,x)$ は「x の単位あたりの増加」がもたらす「将来価値の増分」である。

言い換えれば $\lambda(t)$ は『x が時刻 t で 1 だけ増えると、将来どれだけ得をするか』を表す。

$\lambda(t)$ は x の「隠れた」価値を表すもので、経済学では限界効用 (marginal utility) もしくはシャドウ・プライス (shadow price) と呼ばれる。

Hamiltonian の直観的解釈

x の増分は $\dot{x} = f$ を通じて f でコントロールされるから、Hamiltonian H は以下のように直観的な解釈ができる。

$$H = \underbrace{L}_{\text{即時利益}} + \underbrace{\lambda f}_{\text{将来利益}}$$

ここで（将来利益）は

$\lambda =$（x が単位量あたり増加することの価値）

$f =$（x の増分*）

の積で構成される。つまり最大原理における $\max_u H$ は、「即時利益と将来利益の和を最大化するような u を各時間ごとに選べ」と言っている。

*もっと厳密に言えば、微小時間 Δt で得られる総利益は $H\Delta t = L\Delta t + \lambda f \Delta t$ であり、このうち $f\Delta t$ が x の実際の増分であり、係数 λ をかけることで将来利益に換算される。

生物学の例

最適植林スケジュール

森林を管理し生態系サービスを享受したいとする。

時刻 t における植林量を $u(t)$ (≥ 0) とした時、森林面積 $x(t)$ は

$$\dot{x}(t) = u(t) - dx(t) \qquad (d：死亡率、初期値\ x(0) = x_0)$$

に従い変化する。この時、目的関数

$$J = \int_0^T \Big(\underbrace{x(t)}_{\substack{\text{生態系}\\\text{サービス}}} - \underbrace{c\frac{u(t)^2}{2}}_{\text{植林コスト}} \Big)\mathrm{d}t \quad \longrightarrow \max$$

を最大にする植林スケジュール $u : [0, T] \to \mathbb{R}_{\geq 0}$ を求めよ。

解法

$f(t, x, u) = u - dx,\ L(t, x, u) = x - c\frac{u^2}{2},\ K(x) = 0$ である。

Hamiltonian を以下で定める。

$$H = L + \lambda f = \left(x - c\frac{u^2}{2}\right) + \lambda(u - dx)$$

解くべき問題は

$$\begin{cases} \dot{x} = u - dx, & x(0) = x_0 \quad (x \text{ の従う式}) \\ \dot{\lambda} = -(1 - d\lambda), & \lambda(T) = 0 \quad (\lambda \text{ の従う式}) \\ u \in \operatorname*{argmax}_u \left[x - c\frac{u^2}{2} + \lambda(u - dx)\right] & (u \text{ は常に最適}) \end{cases}$$

解法

$$\dot{\lambda} = -(1 - d\lambda), \qquad \lambda(T) = 0 \quad \text{（λ の従う式）}$$

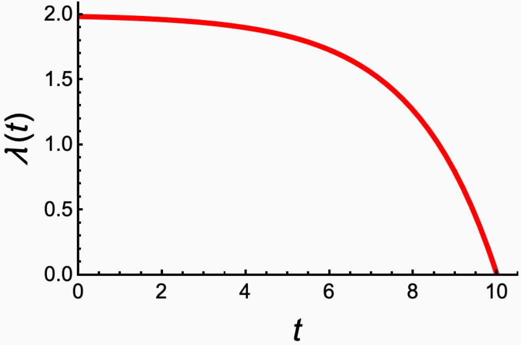
上の条件より

$$\lambda(t) = \frac{1}{d}(1 - e^{d(t-T)})$$

である。

最初のうちは植林し x を増やすことの価値 λ は高い。

終端時刻 T が近づくと、この価値は急速に減少する（将来が少なくなるため）。



解法

$$u \in \underset{u}{\operatorname{argmax}} \left[x - c\frac{u^2}{2} + \lambda(u - dx) \right] \qquad (u \text{ は常に最適})$$

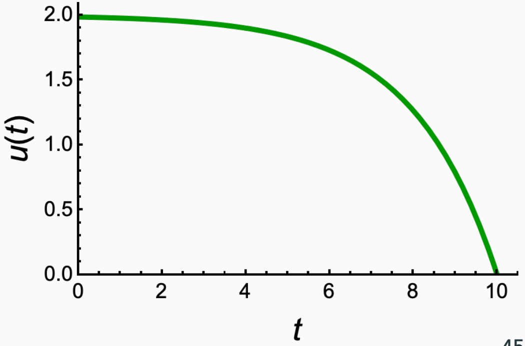
右辺カッコ内は $u = \frac{\lambda}{c}$ を頂点とする上に凸な放物線。

常に $\lambda(t) \geq 0$ なので、$u = \frac{\lambda}{c}$ は $u \geq 0$ を確かに満たす。

よって

$$u(t) = \frac{\lambda(t)}{c} = \frac{1}{cd}(1 - e^{d(t-T)})$$

である。



解法

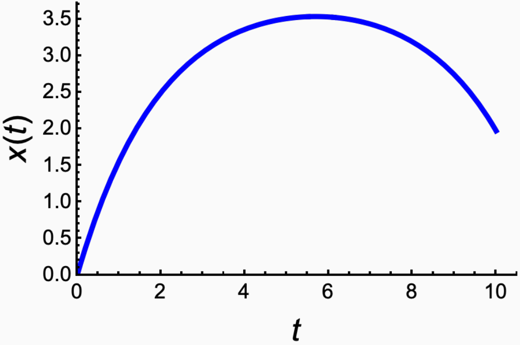
$$\dot{x}(t) = u - dx, \qquad x(0) = x_0 \quad （x\text{ の従う式}）$$

前のスライドの u を代入して

$$\dot{x} = \frac{1}{cd}(1 - e^{d(t-T)}) - dx, \quad x(0) = x_0$$

これを解いて

$$x(t) = x_0 e^{-dt} + \frac{1}{2cd^2}(1 - e^{-dt})\left\{2 - e^{-dT}(1 + e^{dt})\right\}$$

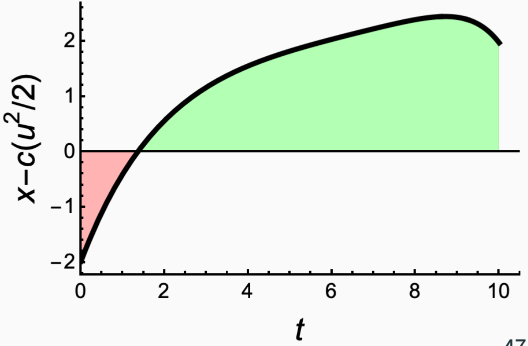


解法

$$J = \int_0^T \left(x(t) - c\frac{u(t)^2}{2} \right) \mathrm{d}t \quad \longrightarrow \max$$

上で求めた $x(t), u(t)$ を代入すると、目的関数の値は

$$J = \frac{1 - e^{-dT}}{d} x_0 + \frac{T}{2cd^2} - \frac{(1 - e^{-dT})(3 - e^{-dT})}{4cd^3}$$



最初は損をしてまで植林する。

Remark

最大原理は必要条件だったので、以上の解析は解の候補を与えるにすぎない。

しかし（！）、今回の場合、先程求めた J が HJB 方程式の解の候補を与えてくれている。

$$V(t,x) = \frac{1-e^{-d(T-t)}}{d}x + \frac{T-t}{2cd^2} - \frac{(1-e^{-d(T-t)})(3-e^{-d(T-t)})}{4cd^3}$$

直接計算により、この V は実際に HJB 方程式の解であることが分かるので、以上の解析で求めた $u(t)$ は実際に最適解である（！）

（HJB 方程式だけを見て、上の解を見つけるのは極めて困難であろう）。

最適解

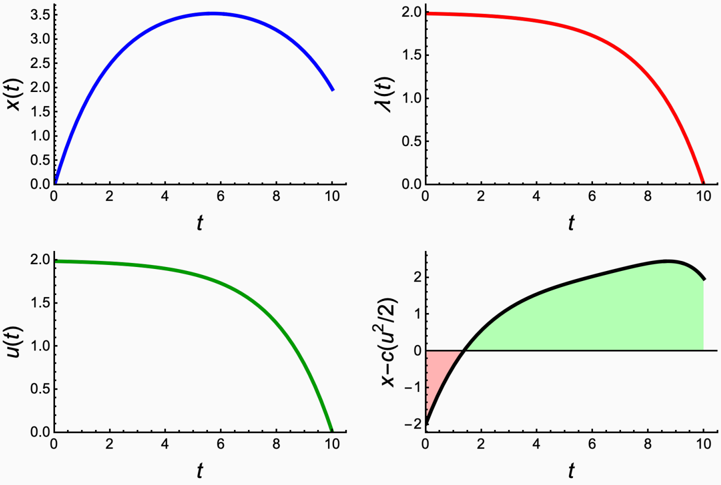
パラメータ：

$x_0 = 0,$

$c = 1,$

$d = 0.5,$

$T = 10$



最適解

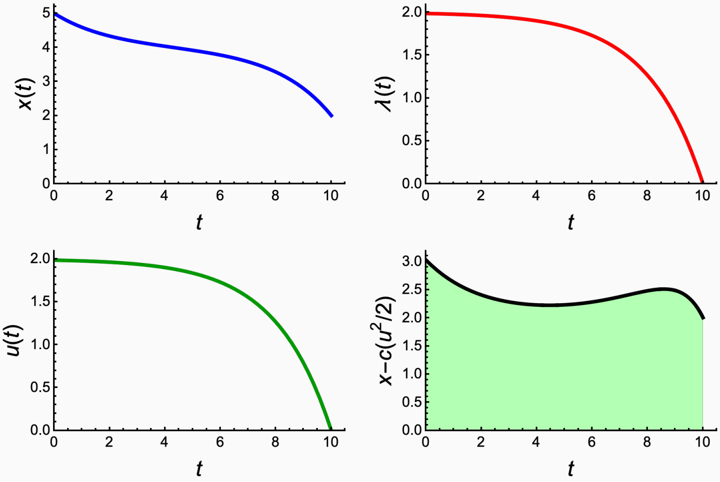
パラメータ：

$x_0 = 5,$

$c = 1,$

$d = 0.5,$

$T = 10$



まとめ

まとめ

- 動的最適化の主要な方法論二つを解説した。
  - 動的計画法では価値関数 V に基づき、Bellman 方程式を立式する。
  - Pontryagin の最大原理では共役状態変数 λ と Hamiltonian H を構成して、微分方程式系を導く。
- いずれも生物学のスケジュールの問題に広く応用できる。

参考文献

- Liberzon, D. (2011). "*Calculus of Variations and Optimal Control Theory: A Concise Introduction.*" Princeton University Press.
- Sethi, S. P. (2021). "*Optimal Control Theory: Applications to Management Science and Economics.*" Springer.