変分法

… ある1つの停留曲線を考えて2点境界値問題を導く.
⇝ 最適制御問題に適用して得られる解は時刻のみの関数.

さまざまな初期状態に対する最適制御全体を考えれば、状態の関数として最適制御入力を表現できる.
⇝ 動的計画法

6章の内容

- 連続時間システムの最適制御問題に動的計画法を適用し、ハミルトン・ヤコビ・ベルマン方程式を導出する.
- ハミルトン・ヤコビ・ベルマン方程式を用いて オイラー・ラグランジュ方程式を一般化した最小原理と呼ばれる条件を導出する.

6.1. ハミルトン・ヤコビ・ベルマン方程式

Given

- 状態方程式

$$\dot{x}(t) = f(x(t), u(t), t) \qquad (6.1)$$

- 初期時刻 t_0, 終端時刻 t_f, 初期状態 x_0

評価関数

$$J = \phi(x(t_f)) + \int_{t_0}^{t_f} L(x(t), u(t), t)\, dt \qquad (6.2)$$

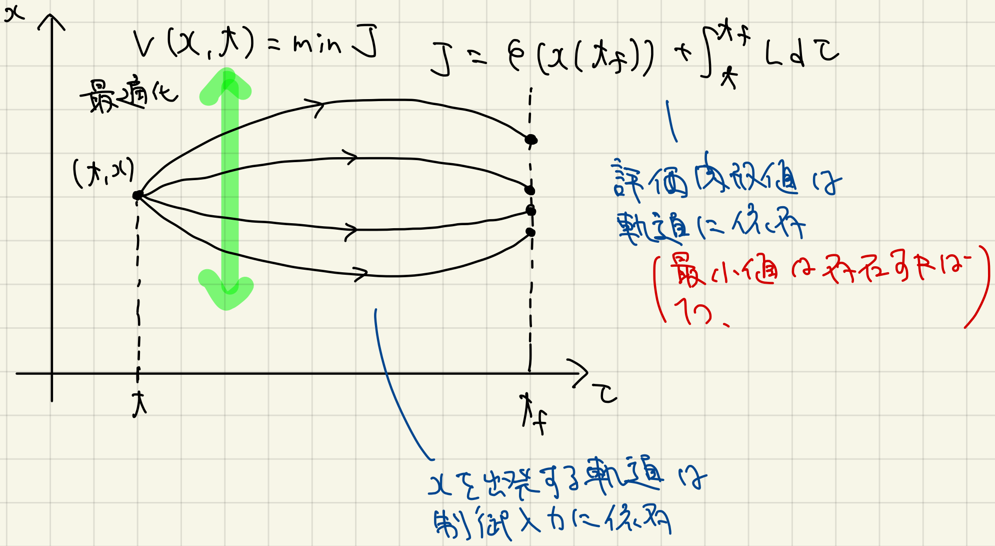
を最小化する最適制御問題.

この最適制御問題を含む一般的な問題として、以下の問題を考える.

ある時刻 t ($t_0 \le t \le t_f$) に、ある状態 x を出発して、終端時刻 t_f までの評価関数

$$J = \phi(x(t_f)) + \int_{t}^{t_f} L(x(\tau), u(\tau), \tau)\, d\tau \qquad (6.3)$$

を最小化する最適制御問題.



最適軌道の評価関数値は、$V:(x,t)\mapsto V(x,t)$ とみなせる。

$(x,t)\in\mathbb{R}^n\times[t_0,t_f]$, $V(x,t)\in\mathbb{R}$

時間区間 $[t,t_f]$ の制御入力関数を $u[t_0,t_f]$ で表すと、関数 $V(x,t)$ は

$$V(x,t)=\min_{u[t,t_f]}\left(\Phi(x(t_f))+\int_t^{t_f}L(x(\tau),u(\tau),\tau)\,d\tau\right)$$

で定義され、値関数と呼ばれる。（※ inf でテイギ？）

- 値関数で $t=t_0$, $x=x_0$ としたときの値 $V(x_0,t_0)$ が、元の最適制御問題における評価関数 (6.2) の最小値。

- $t=t_f$ のとき、積分区間の長さが0であり、$x(t_f)=x$ である。

$$V(x,t_f)=\Phi(x) \qquad (6.4)$$

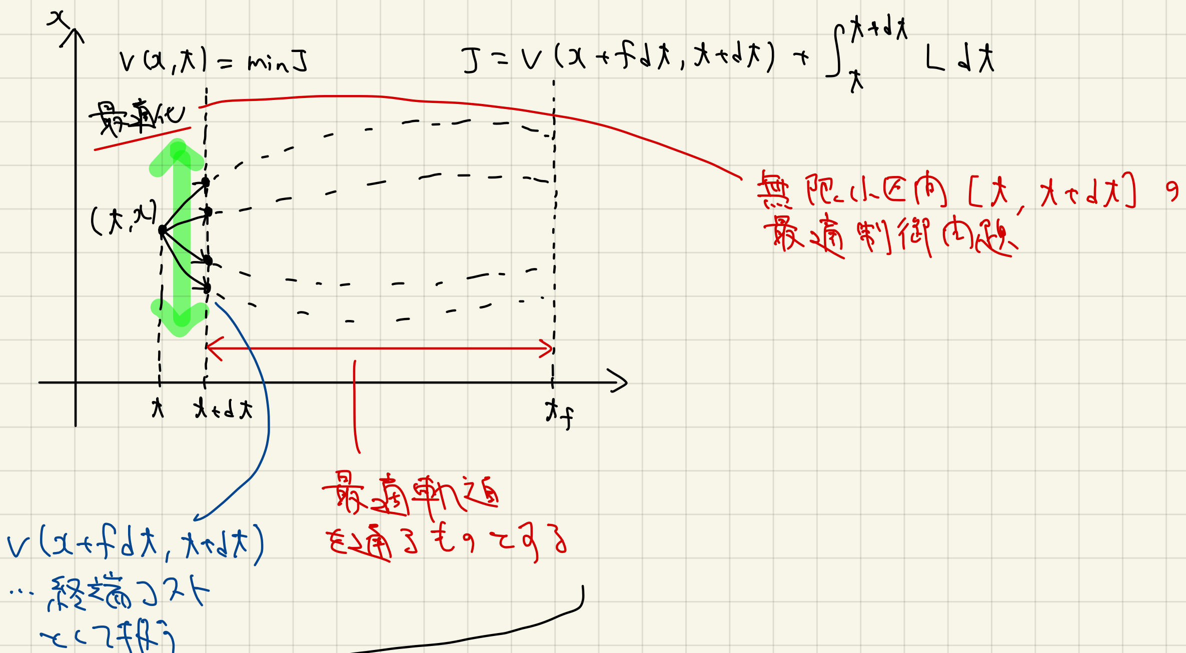
☒ $V(x,t)$ がみたす方程式の導出。

dt：無限小時間として、$V(x,t)$ の定義式を区間 $[t,t+dt]$ と $[t+dt,t_f]$ に分割。

$$V(x,t)=\min_{u[t,t_f]}\left(\int_t^{t+dt}L(x(\tau),u(\tau),\tau)\,d\tau + \Phi(x(t_f))+\int_{t+dt}^{t_f}L(x(\tau),u(\tau),\tau)\,d\tau\right)$$

- 第1項は $u[t+dt,t_f]$ に依存しない
  ∵ $t+dt$ 以降の部分のみが $u[t+dt,t_f]$ によって最小化される。

- 時刻 $t+dt$ に状態は $x+f(x,u,t)dt$ になっている。
  ⇝ $t+dt$ 以降の評価関数の最小値は、
  $V(x+f(x,u,t)dt,\ t+dt)$



このようにしても、最小値は $V(x,t)$ のまま

2次以上のギロンを数式で表現する。

$$V(x,t) = \underbrace{\min_{u[t,t+dt]}}_{\text{微小区間の最適化}} \left\{ \int_t^{t+dt} L(x(\tau), u(\tau), \tau)\, d\tau + \underbrace{\min_{u[t+dt,t_f]} \left(\phi(x(t_f)) + \int_{t+dt}^{t_f} L(x(\tau), u(\tau), \tau)\, d\tau \right)}_{V(x+f(x,u,t)\,dt,\ t+dt)} \right\}$$

$$= \min_{u[t,t+dt]} \left\{ \int_t^{t+dt} L(x(\tau), u(\tau), \tau)\, d\tau + V(x+f(x,u,t)dt, t+dt) \right\}$$

$$= \min_{u} \left(L(x,u,t)\, dt + \underline{V(x+f(x,u,t)\,dt,\ t+dt)} \right) \qquad (6.5)$$

dtが無限小
〜) $u[t, t+dt]$による最小化を、
時刻tでの関数値$u(t)$に相当する
ベクトルuによる最小化におきかえる

$V(x+f\,dt, t+dt)$をTaylor展開（1次近似）

$$V(x+f\,dt, t+dt) = V(x,t) + \begin{bmatrix} \frac{\partial V}{\partial x} & \frac{\partial V}{\partial t} \end{bmatrix} \begin{bmatrix} f\,dt \\ dt \end{bmatrix}$$

$$= V(x,t) + \frac{\partial V}{\partial x}(x,t) f(x,u,t)\, dt + \frac{\partial V}{\partial t}(x,t)\, dt \qquad (*)$$

$(*)$を(6.5)に代入。

$$V(x,t) = \min_u \left(L(x,u,t)dt + V(x,t) + \frac{\partial V}{\partial x}(x,t) f(x,u,t)\, dt + \frac{\partial V}{\partial t}(x,t)\, dt \right)$$

$$\therefore \min_u \left(\underline{L(x,u,t) + \frac{\partial V}{\partial x}(x,t) f(x,u,t)} + \frac{\partial V}{\partial t}(x,t) \right) = 0 \qquad (**)$$

ここで、5.1節で導入したハミルトン関数

$$H(x,u,\lambda,t) = L(x,u,t) + \underline{\lambda^T} f(x,u,t)$$

を用いると、

$$L(x,u,t) + \underline{\frac{\partial V}{\partial x}(x,t)} f(x,u,t) = H\left(x, u, \left(\frac{\partial V}{\partial x}\right)^T(x,t), t\right) \qquad (***)$$

である。さらに、$\frac{\partial V}{\partial t}(x,t)$は$u$に依存しないので、$(**)$, $(***)$から

$$-\frac{\partial V}{\partial t}(x,t) = \min_u H\left(x, u, \left(\frac{\partial V}{\partial x}\right)^T(x,t), t\right) \qquad (6.6)$$

を得る。この方程式は、ハミルトン・ヤコビ・ベルマン方程式と呼ばれる。

@右辺の最小値が存在するとき、$V(x,t)$の偏微分方程式となる。

Point
- 最適制御の評価関数値を初期状態の関数とみなしたところ
- 区間 $[t, t_f]$ の値関数を、部分区間 $[t+dt, t_f]$ の値関数によって再帰的に表しているところ

〜) 動的計画法

以上の議論をまとめると、次の定理を得る。

Th 6.1. (最適制御の必要条件)

> 任意の時刻 $t\ (t_0 \leq t \leq t_f)$ に対し、任意の状態 $x \in \mathbb{R}^n$ から出発し、評価関数 (6.3) を最小化する最適制御問題の解が存在するとする。
> その値関数 $V(x, t)$ が微分可能なとき、終端条件 (6.4) の下で
> ハミルトン・ヤコビ・ベルマン方程式 (6.6) が成立する。
> また、最適制御は (6.6) 右辺の最小値を達成している。

逆に、ハミルトン・ヤコビ・ベルマン方程式が解ければ、最適制御と値関数が得られることも示せる。

ハミルトン・ヤコビ・ベルマン方程式 (6.6) の右辺を最小化する入力 u を $u_{opt}(x, t)$ とすると、任意の u に対して

$$H\left(x, u, \left(\frac{\partial V}{\partial x}\right)^T(x, t), t\right) \geq H\left(x, u_{opt}(x, t), \left(\frac{\partial V}{\partial x}\right)^T(x, t), t\right) = -\frac{\partial V}{\partial t}(x, t)$$

が成立。
等号成立条件は $u = u_{opt}(x, t)$。

ハミルトン関数の定義より、次式を得る。

$$L(x, u, t) \geq -\frac{\partial V}{\partial x} f(x, u, t) - \frac{\partial V}{\partial t}(x, t) = -\dot{V}(x, t)$$

したがって、終端条件 (6.4) に注意して評価関数を計算すると、

$$\begin{aligned}
J &= \underbrace{\phi(x(t_f))}_{V(x(t_f), t_f)} + \int_{t_0}^{t_f} L(x(t), u(t), t)\, dt \\
&\geq V(x(t_f), t_f) + \int_{t_0}^{t_f} \left(-\dot{V}(x, t)\right) dt \\
&= V(x_0, t_0)
\end{aligned}$$

となる。よって、等号成立条件は、すべての時刻で $u(t) = u_{opt}(x(t), t)$ が成り立つことである。

この議論を定理の形でまとめる。

Th 6.2 (最適制御の十分条件)

終端条件 (6.4) の下で ハミルトン・ヤコビ・ベルマン方程式 (6.6) の解 $V(x,t)$ が存在して微分可能だとし、各 x と t に対して (6.6) の右辺の最小値を達成する制御入力 $u = u_{opt}(x,t)$ が存在するとする。また、制御入力を $u = u_{opt}(x,t)$ で与えた閉ループシステム

$$\dot{\bar{x}}(t) = f(\bar{x}(t), u_{opt}(\bar{x}(t), t), t), \quad \bar{x}(t_0) = x_0$$

の解 $\bar{x}(t)$ が存在するとする。このとき、$u_{opt}(\bar{x}(t), t)$ は時刻 t_0 で状態 x_0 から出発する評価関数を最小にする最適制御であり、評価関数の最小値は $V(x_0, t_0)$ である。

Remark

- 制御入力に拘束条件が課せられ、ある集合 $\Omega \subset \mathbb{R}^m$ から制御入力を選ばなければならないときでも、上の定理は成立する。
- 状態に対する拘束条件が評価区間全体や終端時刻において課せられているときは、そのような拘束条件のもとで評価関数がとりうる最小値として値関数を定義する。
- (6.6) の右辺の最小値を達成する制御入力 $u_{opt}(t)$ は一般に陽に求められるとは限らない
  … 入力に関して状態方程式が1次式、評価関数が2次式で入力に拘束条件がない場合には陽に求められる。

例 6.1 入力アフィンシステム。

制御対象と評価関数を少し限定して、

$$\dot{x} = \underbrace{f(x) + B(x)u}_{f(x,u,t)} \tag{6.7}$$

$$J = \phi(x(t_f)) + \int_{t_0}^{t_f} \left(\underbrace{g(x(t)) + \frac{1}{2}u^T R u}_{L(x,u,t)} \right) dt \tag{6.8}$$

とした場合を考える。

@ $f(x) \in \mathbb{R}^n$, $B(x) \in \mathbb{R}^{n\times m}$, $g(x) \in \mathbb{R}$, $R \in \mathbb{R}^{m\times m}$

(6.7) の形のシステムを入力アフィンシステムという。ハミルトン関数は、

$$H\left(x, u, \left(\frac{\partial V}{\partial x}\right)^T\right) = g(x) + \frac{1}{2}u^T R u + \frac{\partial V}{\partial x}(f(x) + B(x)u)$$

である。
… 入力 u に関して2次
⇝ R が正定ならば、H の最小値を達成する u は一意に存在。

$$\frac{\partial H}{\partial u} = u^T R + \frac{\partial V}{\partial x} B(x) = 0$$

を解いて、

$$u_{opt}(x,t) = -R^{-1} B^T(x) \left(\frac{\partial V}{\partial x}\right)^T (x,t) \qquad (6.9)$$

と求められる。

これがハミルトン・ヤコビ・ベルマン方程式 (6.6) 右辺の最小値を与えるので、

$$-\frac{\partial V}{\partial t} = \min_u H\left(x, u, \left(\frac{\partial V}{\partial x}\right)^T\right)$$

$$= \frac{\partial V}{\partial x} f(x) - \frac{1}{2} \frac{\partial V}{\partial x} B(x) R^{-1} B^T(x) \left(\frac{\partial V}{\partial x}\right)^T + g(x) \qquad (6.10)$$

のように、値関数 $V(x,t)$ の偏微分方程式が得られる。

入力を消去した式 (6.10) をハミルトン・ヤコビ方程式ということも多い。

特に、時間 t に依存しない定常解 $V(x)$ が存在したとすると、

$$\frac{\partial V}{\partial x} f(x) - \frac{1}{2} \frac{\partial V}{\partial x} B(x) R^{-1} B(x)^T \left(\frac{\partial V}{\partial x}\right)^T + g(x) = 0 \qquad (6.11)$$

が成り立つ。これは終端時刻 $t_f \to \infty$ で $\partial V / \partial t \to 0$ となった無限評価区間の問題に相当する。

定常解が求められれば、最適制御 (6.9) も時刻に依存しない状態フィードバック制御になることがわかる。

◎ (6.10) は非線形の偏微分方程式
… 解析的に解くことは多くの場合不可能。

解析的に解ける例：LQ制御。

例 6.2. LQ制御

$$\dot{x} = \underbrace{A(t)x + B(t)u}_{f}, \quad x(t_0) = x_0$$

$$J = \frac{1}{2} x^T(t_f) S_f x(t_f) + \int_{t_0}^{t_f} \underbrace{\frac{1}{2}\left(x^T Q(t) x + u^T R(t) u\right)}_{L} dt$$

ハミルトン関数は、

$$H(x, u, \lambda, t) = \frac{1}{2}(x^T Q x + u^T R u) + \lambda^T (Ax + Bu)$$

ハミルトン・ヤコビ・ベルマン方程式と境界条件は、

$$-\frac{\partial V}{\partial t} = \min_u \left\{ \frac{1}{2}(x^T Q x + u^T R u) + \frac{\partial V}{\partial x}(Ax + Bu) \right\}$$

$$V(x, t_f) = \frac{1}{2} x^T S_f x$$

R が正定ならば、最小値を達成する u_{opt} が一意に存在し、

$$\frac{\partial H}{\partial u} = u^T R + \frac{\partial V}{\partial x} B = 0$$

より、

$$u_{opt} = -R^{-1} B^T \left(\frac{\partial V}{\partial x} \right)^T \qquad (6.12)$$

これをハミルトン・ヤコビ・ベルマン方程式に代入して整理すると、

$$-\frac{\partial V}{\partial t} = \frac{1}{2} x^T Q x + \frac{\partial V}{\partial x} A x \underline{- \frac{1}{2} \frac{\partial V}{\partial x} B R^{-1} B^T \left(\frac{\partial V}{\partial x} \right)^T} \qquad (6.13)$$

$$\boxed{\begin{aligned} &\frac{1}{2} u^T R u + \frac{\partial V}{\partial x} B u \\ &= \left(\frac{1}{2} u^T R + \frac{\partial V}{\partial x} B \right) u \\ &= \left(-\frac{1}{2} \left(\frac{\partial V}{\partial x} \right) B (R^{-1})^T R + \frac{\partial V}{\partial x} B \right) \left(-R^{-1} B^T \left(\frac{\partial V}{\partial x} \right)^T \right) \\ &= -\frac{1}{2} \frac{\partial V}{\partial x} B R^{-1} B^T \left(\frac{\partial V}{\partial x} \right)^T \end{aligned}}$$

(6.13) は例 6.1 の式 (6.10) に相当する。$\partial V / \partial x$ の 2 次の項を含む非線形偏微分方程式だが、境界条件を参考にして

$$\underline{V(x, t) = \frac{1}{2} x^T S(t) x} \qquad (6.14)$$

とおいて、(6.13) をみたすように対称行列 $S(t)$ を決定できる。

$$\frac{\partial V}{\partial t} = \frac{1}{2} x^T \dot{S} x, \quad \frac{\partial V}{\partial x} = x^T S$$

(6.14) を (6.13) に代入。

$$\begin{aligned} -\frac{1}{2} x^T \dot{S} x &= \frac{1}{2} x^T Q x + x^T S A x - \frac{1}{2} x^T S B R^{-1} B^T S x \\ &= \frac{1}{2} x^T (Q + \underline{A^T S + S A} - S B R^{-1} B^T S) x \end{aligned}$$

$$x^T S A x = x^T (A^T S + S A) x / 2 \quad \text{for any } x$$

$$\underline{-\dot{S} = Q + A^T S + S A - S B R^{-1} B^T S, \qquad S(t_f) = S_f} \quad \text{リッカチ微分方程式}$$

とすれば、任意の x と t に対して境界条件も含めてハミルトン・ヤコビ・ベルマン方程式が満たされる。

最適制御は (6.12) に (6.14) を代入。

$$u_{opt}(x, t) = -R^{-1} B S x$$

(6.14) より、$x(t_0) = x_0$ から出発する最適軌道の評価関数は、

$$V(x_0, t_0) = \frac{1}{2} x_0^T S(t_0) x_0 \qquad (6.15)$$

$\rightsquigarrow$ 最適制御と最適軌道の評価関数値は、リッカチ微分方程式の解 $S(t)$ によって表される。