システム制御理論と統計的機械学習

第 6 章：漸近挙動

加嶋 健司

October 1, 2025

京都大学情報学研究科

本章の流れ

6.1 動的計画法
無限時間の最適制御問題を定式化し，動的計画法を用いて解法を紹介

6.2 線形最適制御
無限時間 LQR 問題に対して，解法や安定性について議論

動的計画法

無限時間最適制御

問題 6.1.1 - 無限時間最適制御

遷移確率密度関数 Ψ と初期状態 $x_0 \in \mathrm{rv}(\mathbb{X})$ をもつ確率システム

非負ステージコスト $\ell : \mathbb{X} \times \mathbb{U} \to \mathbb{R}_+$，入力許容集合 $U(\mathrm{x}) \subset \mathbb{U}$，**割引率** $\beta \in (0, 1)$

$u_k \in U(x_k)$ を満たし，コスト関数

$$J(u_:) := \mathbb{E}^u \left[\sum_{k=0}^{\infty} \beta^k \ell(x_k, u_k) \right] \tag{1}$$

を最小化する因果的な制御入力 $u_:$ を求めよ．

状態価値関数（k を引数にもたない）

$$\mathrm{V}^*(\mathrm{x}) := \inf_{u_{0:} \in \mathcal{U}} \mathbb{E}^u_{x_0 = \mathrm{x}} \left[\sum_{k=0}^{\infty} \beta^k \ell(x_k, u_k) \right], \quad \mathrm{x} \in \mathbb{X} \tag{2}$$

極限と期待値（積分）の交換

注意 6.1.2 – 極限と期待値（積分）の交換

$$\mathbb{E}\left[\underbrace{\lim_{\bar{k}\to\infty}\sum_{k=0}^{\bar{k}}\ell(k,x_k,u_k)}_{\text{確率変数}}\right] = \underbrace{\lim_{\bar{k}\to\infty}\mathbb{E}\left[\sum_{k=0}^{\bar{k}}\ell(k,x_k,u_k)\right]}_{\text{数列}} \tag{3}$$

- $\infty - \infty$ の不定形があらわれる場合など，仮定なく成り立つわけではない．
- いくつかの十分条件が知られているが，ℓ の非負性もその一つ

ベルマン方程式

$$\mathrm{V}^*(\mathrm{x}) := \inf_{u_{0:}\in\mathcal{U}} \mathbb{E}^u_{x_0=\mathrm{x}}\left[\sum_{k=0}^{\infty}\beta^k \ell(x_k, u_k)\right], \quad \mathrm{x}\in\mathbb{X}$$

補題 6.1.5 – 状態価値関数と初期状態

$$\inf_{u_:\in\mathcal{U}} J(u_:) = \mathbb{E}\left[\mathrm{V}^*(\mathrm{x}_0)\right] \tag{4}$$

定理 6.1.6 – 状態価値関数とベルマン方程式

状態価値関数 V^* は**ベルマン方程式**の解の一つ（$\mathscr{B}[\cdot]$ を**ベルマン作用素**とよぶ）

$$V = \mathscr{B}[V] := \inf_{\mathrm{u}\in U(\mathrm{x})} (\ell + \beta\mathscr{A}_x[V])(\mathrm{x},\mathrm{u}) \tag{5}$$

- $\mathscr{A}_x[V](\mathrm{x},\mathrm{u}) := \int_{\mathbb{X}} V(\mathrm{x}')\Psi(\mathrm{x}'|\mathrm{x},\mathrm{u})\mathrm{dx}' = \mathbb{E}[V(x_{k+1})|x_k = \mathrm{x},\ u_k = \mathrm{u}]$

ベルマン方程式

定理 6.1.7 – 定理状態価値関数の上界

下に有界な関数 $\hat{V} : \mathbb{X} \to \mathbb{R}_+$ が $\hat{V} \geq \mathscr{B}[\hat{V}]$ を満たすとする．このとき，

1. $\hat{V} \geq \mathrm{V}^*$ が成り立ち，
2. 任意の $\epsilon > 0$ に対して，下界の定義より

$$\hat{V}(\mathrm{x}) + \epsilon \geq (\ell + \beta \mathscr{A}_x[\hat{V}])(\mathrm{x}, \pi(\mathrm{x})),\ \pi(\mathrm{x}) \in U(\mathrm{x}), \quad \forall \mathrm{x} \in \mathbb{X} \tag{6}$$

を満たす $\pi : \mathbb{X} \to \mathbb{U}$ が存在し，確定的状態フィードバック $u_k = \pi(x_k)$ は，

$$J(\pi) := \mathbb{E}^{\pi}\left[\sum_{k=0}^{\infty} \beta^k \ell(x_k, u_k)\right] \leq \mathbb{E}\left[\hat{V}(\mathrm{x}_0)\right] + \frac{\epsilon}{1-\beta} \tag{7}$$

を満たす．

ベルマン方程式

系 6.1.8 – 最適制御則

状態価値関数 V^* に対して，

$$\pi^*(\mathrm{x}) := \operatorname*{argmin}_{\mathrm{u}\in U(\mathrm{x})}(\ell + \beta\mathscr{A}_x[\mathsf{V}^*])(\mathrm{x},\mathrm{u}) = \operatorname*{argmin}_{\mathrm{u}\in U(\mathrm{x})} \mathsf{Q}^*(\mathrm{x},\mathrm{u}) \tag{8}$$

が存在すれば，$u_k = \pi^*(x_k)$ は最適制御則である（$J(\pi^*) = \min_u J(u_:)$）.

- 初期分布非依存かつ確定的状態フィードバックであることに加え，時不変
- 閉ループ安定性を保証（前ページ最終式右辺は有限が有限であることから）

$$\lim_{k\to\infty} \beta^k \mathbb{E}^\pi\left[\ell(x_k, u_k)\right] = 0 \tag{9}$$

ベルマン方程式

定理 6.1.9 – 状態価値関数の下界

上に有界な関数 $\check{\mathrm{V}}: \mathbb{X} \to \mathbb{R}$ が

$$\check{\mathrm{V}}(\mathrm{x}) \leq (\ell + \beta \mathscr{A}_x[\check{\mathrm{V}}])(\mathrm{x}, \mathrm{u}), \quad \forall \mathrm{x} \in \mathbb{X},\ \mathrm{u} \in U(\mathrm{x}) \tag{10}$$

を満たすならば，$\check{\mathrm{V}} \leq \mathrm{V}^*$ が成り立つ．

価値反復法

ベルマン方程式を直接的に解かない，反復法による解法

補題 6.1.10 – ベルマン作用素の単調性

$\mathbb{X}$ で定義された実数値関数 $V_1,\ V_2$ が $V_1 \leq V_2$ を満たすとき，$\mathscr{B}[V_1] \leq \mathscr{B}[V_2]$ が成り立つ．

補題 6.1.11 – 有界な関数に対するベルマン作用素の縮小性

任意の有界な関数 $V_1,\ V_2 : \mathbb{X} \to \mathbb{R}$ に対して，

$$\|\mathscr{B}^i[V_1] - \mathscr{B}^i[V_2]\|_\infty \leq \beta^i \|V_1 - V_2\|_\infty,\ \forall i \in \mathbb{Z}_+ \tag{11}$$

が成り立つ．ただし，$\|V\|_\infty := \sup_{\mathrm{x}\in\mathbb{X}} |V(\mathrm{x})|$ である．

- $\mathscr{B}^{i+1}[V] := \mathscr{B}[\mathscr{B}^i[V]]$

価値反復法

定理 6.1.12 – 価値反復法（value iteration）

状態価値関数 V^* は有界な関数であるとする．このとき，V^* はベルマン方程式 $V = \mathscr{B}[V]$ に対する唯一の有界な解であり，任意の有界な関数 $V_0 : \mathbb{X} \to \mathbb{R}$ に対して，

$$\lim_{i\to\infty} \|\mathscr{B}^i[V_0] - \mathsf{V}^*\|_\infty = 0 \tag{12}$$

が成り立つ．

価値反復法

反復中に得られる関数は有限区間最適制御問題の解

補題 6.1.13 – 終端コストと状態価値関数

任意の関数 $V : \mathbb{X} \to \mathbb{R}_+$ および $\bar{k} \geq 1$ に対して，

$$\mathscr{B}^{\bar{k}}[V](\mathrm{x}) = \inf_{u_{0:\bar{k}-1} \in \mathcal{U}} \mathbb{E}^{u}_{x_0=\mathrm{x}} \left[\sum_{k=0}^{\bar{k}-1} \beta^k \ell(x_k, u_k) + \beta^{\bar{k}} V(x_{\bar{k}}) \right] \tag{13}$$

であり，特に，状態価値関数 V^* および任意の $\bar{k} \geq 0$ に対して，

$$\mathsf{V}^*(\mathrm{x}) = \inf_{u_{0:\bar{k}-1} \in \mathcal{U}} \mathbb{E}^{u}_{x_0=\mathrm{x}} \left[\sum_{k=0}^{\bar{k}-1} \beta^k \ell(x_k, u_k) + \beta^{\bar{k}} \mathsf{V}^*(x_{\bar{k}}) \right] \tag{14}$$

- 後者の性質は，モデル予測制御の安定性保証などにおいても用いられる

線形最適制御

安定性

時間とともに減衰する係数のもとでの線形確率システムの挙動を解析する．

定理 6.2.1 – 割引率とシューア安定性

$\beta \in (0,1)$ とし，確定的な初期状態 x_i をもつ線形システム

$$x_{k+1} = Ax_k + v_k,\ x_0 = \mathsf{x}_i \tag{15}$$

において v_k は時間的に無相関かつ $\mathbb{E}[v_k] = 0,\ \mathrm{Var}[v_k] \preceq N$ とする．このとき，ある $Q \succ O$ に対して

$$\beta^k \mathbb{E}[{x_k}^\top Q x_k] \to 0,\ \forall \mathsf{x}_i \tag{16}$$

ならば $\sqrt{\beta}\rho(A) < 1$ が成り立ち，逆に，$\sqrt{\beta}\rho(A) < 1$ ならば，任意の $Q \succ O$ に対して式 (16) が成り立つ．

安定性

安定性とリアプノフ方程式の解の関係に関しても，つぎの結果が成り立つ．

定理 6.2.2 – 割引率とリアプノフ方程式

線形確率システム

$$x_{k+1} = Ax_k + v_k,\ x_0 = \mathrm{x}_i \tag{17}$$

を考える．このとき，ある $Q \succ O$ に対してリアプノフ方程式

$$\beta A^\top \Pi A + Q = \Pi \tag{18}$$

を満たす $\Pi \succ O$ が存在するならば $\sqrt{\beta}\rho(A) < 1$ である．逆に $\sqrt{\beta}\rho(A) < 1$ ならば，任意の $Q \succeq O$ に対して式 (18) は唯一解 $\bar{\Pi} \succeq Q$ をもち，

$$\mathbb{E}\left[\sum_{k=0}^{\infty} \beta^k x_k^\top Q x_k\right] = \mathrm{x}_i^\top \bar{\Pi} \mathrm{x}_i + \bar{\gamma} \tag{19}$$

を満たす x_i に依存しない $\bar{\gamma} \geq 0$ が存在する．

無限区間線形二次レギュレータ

問題 6.2.3 – 無限区間 LQR

線形確率システムに対する二次コスト最小化

$$
\begin{aligned}
&x_{k+1} = Ax_k + Bu_k + v_k,\ \mathbb{E}[v_k] = 0,\ \mathbb{E}[v_k v_k^\top] = N,\ \mathbb{E}[v_k v_l^\top] = O,\ k \neq l \\
&\ell(\mathrm{x}, \mathrm{u}) := \mathrm{x}^\top Q \mathrm{x} + \mathrm{u}^\top R \mathrm{u},\ Q,\ R \succ O
\end{aligned}
\tag{20}
$$

- 仮定：$\sqrt{\beta}\rho(A - B\bar{K}) < 1$ を満たす $\bar{K}$ が存在する

無限区間線形二次レギュレータ

定理 6.2.4 – 無限時間 LQR 問題の最適解

$$\begin{aligned}\tilde{R}(\Pi) &:= R + \beta B^\top \Pi B, \tilde{S}(\Pi) := \beta A^\top \Pi B,\ \tilde{Q}(\Pi) := Q + \beta A^\top \Pi A \\ \mathrm{Ric}(\Pi) &:= \tilde{Q}(\Pi) - \tilde{S}(\Pi)\tilde{R}(\Pi)^{-1}\tilde{S}(\Pi)^\top,\end{aligned} \tag{21}$$

と表記すると，**リッカチ方程式**（Riccati equation）とよばれる

$$\mathrm{Ric}(\Pi) = \Pi \tag{22}$$

は唯一の正定解 Π^* をもつ．また，線形制御則

$$u_k := -K(\Pi^*)x_k, K(\Pi) := \tilde{R}(\Pi)^{-1}\tilde{S}(\Pi)^\top \tag{23}$$

は最適解であり，$\sqrt{\beta}\rho(A - BK(\Pi^*)) < 1$ が成り立つ．

- $\gamma(\Pi^*) := \frac{\beta}{1-\beta}\mathrm{Trace}(N\Pi^*)$ で定義される $V^*(\mathrm{x}) := \mathrm{x}^\top \Pi^* \mathrm{x} + \gamma(\Pi^*)$ が状態価値関数

価値反復法

$V = x^\top \Pi x + \gamma$ に対して，

$$\mathscr{B}[V] = x^\top \mathrm{Ric}(\Pi)x + \beta(\gamma + \mathrm{Trace}(N\Pi))$$

例 6.2.6 – リカッチイタレーション

不安定な線形システム（$\rho(A) = 1.5571$）

$$A := \begin{bmatrix} 0.8 & 0.9 & 0.86 \\ 0.3 & 0.25 & 1 \\ 0.1 & 0.55 & 0.5 \end{bmatrix},\ B := \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},\ v_k \sim \mathcal{N}(0,1),$$

$$\ell(x,u) := \|x\|^2 + u^2,\ \beta = 0.95$$

$\Pi_0 = O,\ \Pi_{i+1} = \mathrm{Ric}(\Pi_i)$ とすると，$i = 5$ ではじめて安定化制御則（$\sqrt{\beta}\rho(A - BK(\Pi_5)) = 0.9339$）

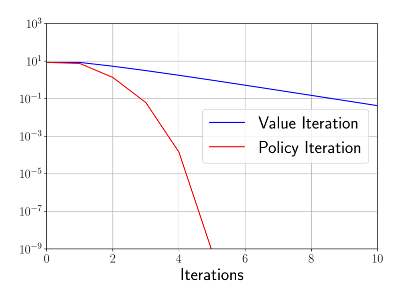

Figure 1: $\Pi_0 = \Pi_{\mathrm{ini}} := \Pi_5$ と初期化した場合の誤差 $\|\Pi_i - \Pi^*\|$
（Policy Iteration は後出）

定常フィルタゲインの安定性

- 無限区間 LQR 問題においても分離定理が成り立つ（$u_k = -K(\Pi^*)\hat{x}_k$ が最適）
- フィードバックゲインは時刻に依存しないのに対して，フィルタゲインは時変であるが，定数行列に収束

定理 6.2.7 – 定常フィルタゲインの安定性

$$\tilde{N}(\Sigma) := N + A\Sigma A^\top,\ \tilde{M}(\Sigma) := M + C\Sigma C^\top,\ \tilde{L}(\Sigma) := A\Sigma C^\top$$
$$\mathrm{Ric_f}(\Sigma) := \tilde{N} - \tilde{L}\tilde{M}^{-1}\tilde{L}^\top,$$

と定義すると，

$$\Sigma_{k+1} = \mathrm{Ric_f}(\Sigma_k),\ H_k = \tilde{L}(\Sigma_k)\tilde{M}(\Sigma_k)^{-1}$$

で与えられる Σ_k はリッカチ方程式

$$\Sigma = \mathrm{Ric_f}(\Sigma)$$

の一意正定解に収束し，H_k の収束値 H に対して $A - HC$ はシューア安定

長時間平均コストと最適 H^2 制御

減衰率 $\beta = 1$ の場合

- 線形確定システム $x_{k+1} = Ax_k + Bu_k$ の場合，定理 6.2.4 はそのまま成り立つ ($\Pi = \mathrm{Ric}(\Pi)$ が一意正定解 Π^* をもち，$u_k := -K(\Pi^*)x_k,\ \gamma(\Pi^*) = 0$)
- 確率システムに拡張すると評価関数が有界とならない．
- 長時間平均コスト

$$J_{\mathrm{st}}(u_:) := \limsup_{k\to\infty} \mathbb{E}^u \left[\frac{1}{k} \sum_{i=0}^{k} \ell(x_i, u_i) \right] \tag{24}$$

長時間平均コストと最適 H^2 制御

定理 6.2.10 – 長時間平均コストに対するベルマン方程式の不等式評価

$\beta = 1$ とした無限時間最適制御問題

- 非負値関数 $\hat{V} : \mathbb{X} \to \mathbb{R}_+$，$\pi : \mathbb{X} \to \mathbb{U}$，$\gamma \geq 0$ が

$$\hat{V}(\mathrm{x}) + \gamma \geq (\ell + \mathscr{A}_x[\hat{V}])(\mathrm{x}, \pi(\mathrm{x})),\ \pi(\mathrm{x}) \in U(\mathrm{x}) \quad \mathrm{x} \in \mathbb{X}, \tag{25}$$

  を満たすとき，$u_k := \pi(x_k)$ は $J_{\mathrm{st}}(u_:) \leq \gamma$ を満たす．
- 逆に，上に有界な関数 $\check{\mathrm{V}} : \mathbb{X} \to \mathbb{R}$ および $\gamma \geq 0$ が

$$\check{\mathrm{V}} + \gamma \leq \mathscr{B}[\check{\mathrm{V}}] \tag{26}$$

  を満たすとき，任意の $u_:$ に対して $J_{\mathrm{st}}(u_:) \geq \gamma$ である．

- 長時間平均コストは定常状態しか評価しておらず，初期分布に依存しない

長時間平均コストと最適 H^2 制御

系 6.2.11 – 長時間平均コストに対する最適制御則

$V + \gamma = \mathscr{B}[V]$ を満たす有界な関数 V および $\gamma \geq 0$ に対して，

$$\pi^*(\mathrm{x}) := \operatorname*{argmin}_{\mathrm{u} \in U(\mathrm{x})} (\ell + \mathscr{A}_x[V])(\mathrm{x}, \mathrm{u}) \tag{27}$$

が存在すれば，$u_k = \pi^*(x_k)$ は最適制御則である．

この結果を線形システムに適用すると，つぎの結果が得られる．

定理 6.2.12 – 長時間平均コストをもつ LQR

定理 6.2.4 の制御則は，長時間平均コストの場合も最適であり，

最適コストは $\mathrm{Trace}(N\Pi^*)$ である．

- 部分観測の場合も同様であり，伝達関数を用いた定式化は**最適 H^2 制御**とよばれる．