システム制御理論と統計的機械学習

第 4 章：最適制御

加嶋 健司

September 29, 2025

京都大学情報学研究科

本章の流れ

4.1 動的計画法
確率システムの最適制御問題を定式化し，動的計画法による解法を解説

4.2 線形最適制御
有限区間線形二次レギュレータ（LQR）問題を紹介

最適制御

因果性とフィードバック制御 (1)：扱うシステム

離散時間の確率システム制御理論では

- 関数 $f : \mathbb{X} \times \mathbb{U} \times \mathbb{V} \to \mathbb{X}$
- 初期状態 $x_0 \in \mathrm{rv}(\mathbb{X})$
- $\mathbb{X}$ 値の状態確率過程 $x_:$
- $\mathbb{U}$ 値の入力確率過程 $u_:$
- $\mathbb{V}$ 値の雑音過程 $v_:$

に対して

$$x_{k+1} = f(x_k, u_k, v_k), k \in \mathbb{Z}_+ \tag{4.1}$$

により定まる確率システム (stochastic system) を扱う

フィードバック制御

例 4.1.1 – 確定システムのフィードバック制御

確率的な要素を含まない

$$x(k+1) = ax(k) + u(k),\ a > 1,\ x(0) = 1$$

- 実現できる $x(k)$ の軌道は同じ
  FF $u(k) := -Kx(k)$
  FB $u(k) := -K(a-K)^k x(0)$
- 非線形時変でも同様

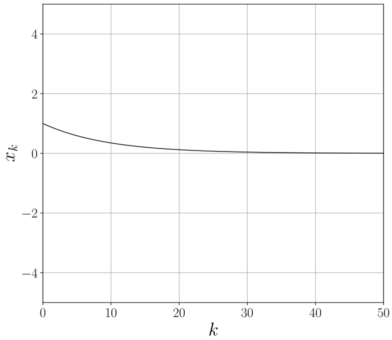


Figure 1: 確定システム ($a = 1.1$)

フィードバック制御

例 4.1.1 – 確率システムのフィードバック制御

確率システム

$$x_{k+1} = ax_k + u_k + 0.1v_k,\ a > 1,\ x_0 = 1$$
$$v_k \sim \mathcal{N}(0,1),\ \text{i.i.d.}$$

- $u_k = -Kx_k$ により $\mathrm{Var}[x_k]$ は発散しない
- 任意の確定的な入力に対して $\mathrm{Var}[x_k] \to \infty$
- 初期状態や状態遷移に不確実性を含む確率システムに対しては，フィードバック制御でしか実現できない挙動

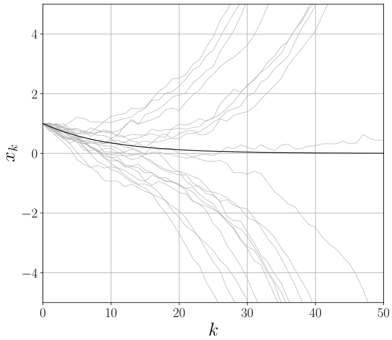


Figure 2: 確率システム ($a = 1.1$)

確率的制御則の例

例 4.1.2 – ディザ量子化器

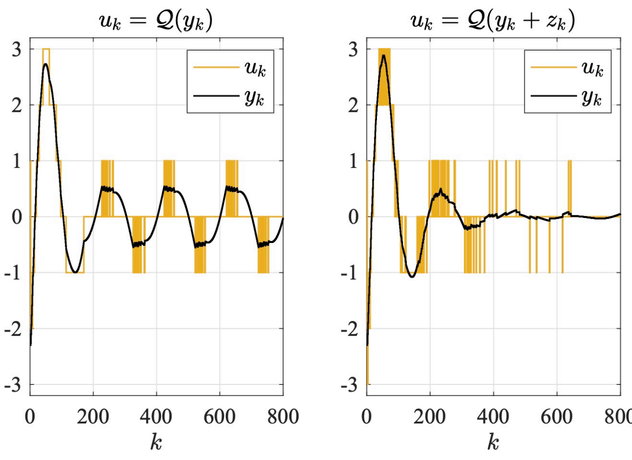
$x_{k+1} = Ax_k + Bu_k,\ \rho(A+BC) < 1$

u_k として整数値しか入力できない

- 確定的制御則 $u_k = y_k$ により $u,\ y$ は速やかに 0 に収束
- 最も近い整数を出力する $u_k = \mathcal{Q}(y_k)$
- 確率的状態フィードバック

$$u_k = \mathcal{Q}(y_k + z_k)$$

$$z_k \sim \textsf{Uni}([-0.5, 0.5]), \text{i.i.d.}$$



因果性とフィードバック制御 (2)：因果性

$x_{k+1} = f_k(x_k, u_k, v_k)$，$v_:$ は無記憶 $(x_0, v_{0:k-1}) \perp\!\!\!\perp v_k$

知り得ない未来の雑音の実現値を利用しない入力

定義 4.1.3 – 入力 $u_:$ の因果性

$$u_{0:k} \perp\!\!\!\perp v_k, \ \forall k \in \mathbb{Z}_+ \tag{4.2}$$

- 任意の因果的な入力 u に対して，$(x_{0:k}, u_{0:k}) \perp\!\!\!\perp v_k$
- **遷移確率密度関数** $\Psi_k(\mathrm{x}_{k+1}|\mathrm{x}_k, \mathrm{u}_k) := \varphi(x_{k+1}|x_k, u_k)$ は，関数 f と v_k の分布で定まる

$$\varphi(\mathrm{x}_{k+1}|\mathrm{x}_{0:k}, \mathrm{u}_{0:k}) = \Psi_k(\mathrm{x}_{k+1}|\mathrm{x}_k, \mathrm{u}_k), \ \forall k, \ \mathrm{x}_{0:k}, \ \mathrm{u}_{0:k} \tag{4.3}$$

因果性とフィードバック制御 (3)：遷移確率密度関数

Example

$x_{k+1} = f'(x_k, u_k) + g(x_k, u_k)v_k,\ v_k \sim \mathcal{N}(0, I)$, i.i.d. の遷移確率密度関数は，

$$\Psi_k(\mathrm{x}'|\mathrm{x}, \mathrm{u}) = \mathcal{N}(\mathrm{x}'|f'(\mathrm{x}, \mathrm{u}), g(\mathrm{x}, \mathrm{u})g(\mathrm{x}, \mathrm{u})^\top)$$

- 時間発展

$$\varphi(\mathrm{x}_{k+1}) = \int_{\mathbb{X}} \int_{\mathbb{U}} \Psi_k(\mathrm{x}_{k+1}|\mathrm{x}_k, \mathrm{u}_k)\varphi(\mathrm{x}_k, \mathrm{u}_k)\mathrm{d}\mathrm{u}_k \mathrm{d}\mathrm{x}_k$$

- 状態入力系列

$$\varphi(\mathrm{x}_{0:k}, \mathrm{u}_{0:k-1}) = \varphi(x_0) \prod_{i=0}^{k-1} \left\{ \Psi_i(\mathrm{x}_{i+1}|\mathrm{x}_i, \mathrm{u}_i)\varphi(u_i|\mathrm{x}_{0:i}, \mathrm{u}_{0:i-1}) \right\}$$

- 因果的な入力 u_k は一般性を失わず過去の状態 $x_{0:k}$ と入力 $u_{0:k-1}$ の実現値により定まる確率分布にしたがうとしてよい

確率システムの最適制御問題 (1)：定式化

問題 4.1.5 – 有限時間最適制御

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

終端時刻 $\bar{k}$，ステージコスト $\ell : \mathbb{X} \times \mathbb{U} \to \mathbb{R}$，終端コスト $\ell_f : \mathbb{X} \to \mathbb{R}$

許容入力集合 $U(\mathrm{x}) \subset \mathbb{U}$

$u_k \in U(x_k)$ を満たし，コスト関数

$$J(u_:) := \mathbb{E}^u \left[\sum_{k=0}^{\bar{k}-1} \ell(x_k, u_k) + \ell_f(x_{\bar{k}}) \right] \tag{4.4}$$

を最小化する因果的な制御入力 $u_:$ を求めよ．

確率系の最適制御問題 (2)：確率的状態フィードバックの十分性

定理 4.1.6 – 確率的状態フィードバック則による因果的制御則の複製

遷移密度関数 Ψ_k および初期状態 $x_0 \in \mathrm{rv}(\mathbb{X})$ をもつ確率システム

任意の因果的な入力 $u^{\mathrm{c}}_{:}$ に対して，対応する状態過程を $x^{\mathrm{c}}_{:}$ とおき，

$$\pi^{\mathrm{s}}_k(\mathrm{u}|\mathrm{x}) := \varphi_{u^{\mathrm{c}}_k}(\mathrm{u}|x^{\mathrm{c}}_k = \mathrm{x}), \quad \forall \mathrm{u} \in \mathbb{U} \tag{4.5}$$

このとき，確率的状態フィードバック則 $\pi^{\mathrm{s}}_k(\cdot|\cdot)$ のもとでの状態・入力過程を $x^{\mathrm{s}}_{:}$, $u^{\mathrm{s}}_{:}$ とおくと，任意の $k \in \mathbb{Z}_+$ に対して $(x^{\mathrm{c}}_k, u^{\mathrm{c}}_k)$ と $(x^{\mathrm{s}}_k, u^{\mathrm{s}}_k)$ は同分布をもつ．

- 各時刻での分布 $\{\varphi_{(x_k, u_k)}\}_{k=0,1,\cdots,\bar{k}}$ により値が定まる評価関数の場合，履歴非依存の確率的状態フィードバックで十分

因果性とフィードバック制御 (6)：状態フィードバックの表現

制御則が過去の履歴を用いない（静的（static））とき，状態 $x_{:}$ はマルコフ過程

- φ_{u_k} が x_k の実現値のみに依存する**確率的状態フィードバック**が

$$\varphi(\mathrm{u}_k|\mathrm{x}_k) = \pi_k(\mathrm{u}_k|\mathrm{x}_k)$$

で与えられるとき，遷移確率密度関数は

$$\varphi(x_{k+1}|\mathrm{x}_k) = \int_{\mathbb{U}} \Psi_k(\mathrm{x}_{k+1}|\mathrm{x}_k, \mathrm{u}_k)\pi_k(\mathrm{u}_k|\mathrm{x}_k)\mathrm{du}_k$$

- $u_k \in \mathrm{rv}(x_k)$ である**確定的状態フィードバック**が

$$u_k := \pi_k(x_k)$$

で与えられるとき，遷移確率密度関数は

$$\varphi(x_{k+1}|\mathrm{x}_k) = \Psi_k(\mathrm{x}_{k+1}|\mathrm{x}_k, \pi_k(\mathrm{x}_k))$$

ベルマン方程式 (1) 状態価値関数

$x_k = \mathrm{x}$ の時に達成できるその後の最適累積コスト

Definition (状態価値関数)

確率最適制御問題に対する**状態価値関数**（state value function）

$$\mathsf{V}^*(k, \mathrm{x}) := \inf_{u_{k:} \in \mathcal{U}} \underbrace{\mathbb{E}^u_{x_k = \mathrm{x}} \left[\sum_{i=k}^{\bar{k}} \ell(i, x_i, u_i) \right]}_{=: \mathsf{V}^u(k, \mathrm{x})} \tag{4.6}$$

補題 4.1.8 – 状態価値関数と初期状態

$$\inf_{u_: \in \mathcal{U}} J(u_:) = \mathbb{E}[\mathsf{V}^*(0, x_0)] \tag{4.7}$$

- x_0 の実現値によって制御則を変えることは因果性に反さないため

ベルマン方程式 (1) 最適性原理

- **最適性原理**（principle of optimality）：最適経路の部分列も最適である
- 確定システム $x(k+1) = f(x(k), u(k))$ の場合

$$\mathrm{V}^*(k, \mathrm{x}) = \inf_{\mathrm{u} \in U(\mathrm{x})} \{\ell(\mathrm{x}, \mathrm{u}) + \mathrm{V}^*(k+1, f(\mathrm{x}, \mathrm{u}))\} \tag{4.8}$$

定理 4.1.9 – 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**（Bellman equation）の唯一解

$$V(\bar{k}, \mathrm{x}) = \ell_f(\mathrm{x}), \ \forall \mathrm{x} \in \mathbb{X}, \tag{4.9a}$$

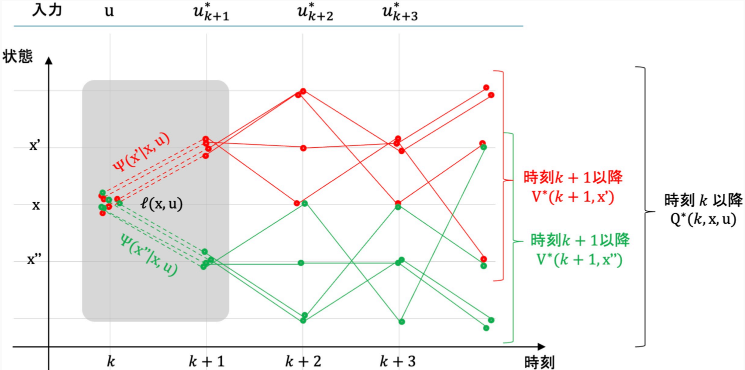
$$V(k, \mathrm{x}) = \inf_{\mathrm{u} \in U_k(\mathrm{x})} (\ell + \mathscr{A}_x[V])(k, \mathrm{x}, \mathrm{u}), \ \forall (k, \mathrm{x}) \in \lfloor \bar{k} \rfloor \times \mathbb{X} \tag{4.9b}$$

$$\mathcal{A}_x[V](k, \mathbf{x}, \mathbf{u}) := \mathbb{E}[V(k+1, x_{k+1}) | x_k = \mathbf{x}, u_k = \mathbf{u}] = \int_{\mathbb{X}} V(k+1, \mathbf{x}')\Psi_k(\mathbf{x}'|\mathbf{x}, \mathbf{u}) d\mathbf{x}'$$

動的計画法 (3)：行動価値関数

注意 4.1.10 – 行動価値関数（action value function）もしくは Q 関数

$$\mathrm{Q}^*(k, \mathrm{x}, \mathrm{u}) := \inf_{u_{k+1:}} \left\{ \ell(k, \mathrm{x}, \mathrm{u}) + \mathbb{E}^{u_{k+1:}}_{x_k=\mathrm{x}, u_k=\mathrm{u}} \left[\sum_{i=k+1}^{\bar{k}} \ell(i, x_i, u_i) \right] \right\}$$



動的計画法 (4)：状態価値関数の上界

定理 4.1.11 – 状態価値関数の上界

$$\begin{aligned}&\hat{V}(\bar{k},\mathrm{x}) \geq \ell_f(\mathrm{x}),\ \forall \mathrm{x} \in \mathbb{X},\\ &\hat{V}(k,\mathrm{x}) \geq \inf_{\mathrm{u}\in U_k(\mathrm{x})} (\ell + \mathscr{A}_x[\hat{V}])(k,\mathrm{x},\mathrm{u}),\ \forall k \in \lfloor \bar{k} \rfloor, \forall \mathrm{x} \in \mathbb{X}\end{aligned} \tag{4.10}$$

を満たす任意の $\hat{V} : [\bar{k}] \times \mathbb{X} \to \mathbb{R}$ に対して，

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

$$\hat{V}(k,\mathrm{x}) + \varepsilon \geq (\ell + \mathscr{A}_x[\hat{V}])(k,\mathrm{x},\pi_k(\mathrm{x})),\ \pi_k(\mathrm{x}) \in U(\mathrm{x}),\ \forall k \in \lfloor \bar{k} \rfloor, \forall \mathrm{x} \in \mathbb{X},$$

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

$$J(\pi) := \mathbb{E}^\pi\left[\sum_{k=0}^{\bar{k}} \ell(k,x_k,u_k)\right] \leq \mathbb{E}\left[\hat{V}(0,x_0)\right] + \epsilon\bar{k}$$

動的計画法 (4)：最適制御則

特に，状態価値関数 V^* が与えられるとき，ベルマン方程式から V^* は (4.10) の等式を満たすことから，$\epsilon = 0$ ととると最小化解はつぎのように求まる．

系 4.1.12 – 最適制御則

状態価値関数 V^* に対して，

$$\pi_k^*(\mathrm{x}) := \operatorname*{argmin}_{\mathrm{u}\in U_k(\mathrm{x})}(\ell + \mathscr{A}_x[\mathsf{V}^*])(k, \mathrm{x}, \mathrm{u}) (= \operatorname*{argmin}_{\mathrm{u}\in U_k(\mathrm{x})} \mathsf{Q}^*(k, \mathrm{x}, \mathrm{u})) \tag{4.11}$$

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

動的計画法 (4)：状態価値関数の下限

定理 4.1.13 – 状態価値関数の下限

関数 $\check{\mathrm{V}} : [0, \bar{k}-1] \times \mathbb{X} \to \mathbb{R}_+$ が

$$\check{\mathrm{V}}(\bar{k}, \mathrm{x}) \le \ell_f(\mathrm{x}), \ \forall \mathrm{x} \in \mathbb{X}, \tag{4.12a}$$

$$\check{\mathrm{V}}(k, \mathrm{x}) \le (\ell + \mathscr{A}_x[\check{\mathrm{V}}])(k, \mathrm{x}, \mathrm{u}), \ \forall k \in \lfloor \bar{k} \rfloor, \ \mathrm{x} \in \mathbb{X}, \ \mathrm{u} \in U_k(\mathrm{x}) \tag{4.12b}$$

を満たすならば，$\check{\mathrm{V}} \le \mathrm{V}^*$ である．

随伴法 (1)：確定離散時間システムにおける最適制御問題

確定離散時間システム

$$x_{k+1} = f_k(x_k, u_k), k \in [0, \bar{k} - 1] \tag{4.13}$$

におけるコスト関数

$$J(u_:) := \sum_{k=0}^{\bar{k}-1} l_k(x_k, u_k) + l_f(x_{\bar{k}}) \tag{4.14}$$

を最小化する制御入力を設計する最適制御問題を考える．この問題は等式拘束条件(4.13)を満たす決定変数 $x_{1:\bar{k}}, u_{0:\bar{k}-1}$ をもつ最適化問題に過ぎない．従って，ラグランジュの未定乗数法を適用できる．

随伴法 (2)：確定離散時間システムに対する未定乗数法

$$\min J(u_:) := \sum_{k=0}^{\bar{k}-1} l_k(x_k, u_k) + l_f(x_{\bar{k}}) \quad \text{s.t.} \quad x_{k+1} = f_k(x_k, u_k) \tag{4.15}$$

ラグランジュの未定乗数 $\zeta_k \in \mathbb{R}^{n_x}$ を導入

$$L(\mathrm{x}_{1:\bar{k}}, \mathrm{u}_{0:\bar{k}-1}, \zeta_{0:\bar{k}-1}) := \ell_f(\mathrm{x}_{\bar{k}}) + \sum_{k=0}^{\bar{k}-1} \left\{ \ell(k, \mathrm{x}_k, \mathrm{u}_k) - \zeta_k^\top (\mathrm{x}_{k+1} - f_k(\mathrm{x}_k, \mathrm{u}_k)) \right\} \tag{4.16}$$

L の停留条件

$$\sum_{k=1}^{\bar{k}} \frac{\partial L}{\partial \mathrm{x}_k}(x_{1:\bar{k}}, u_{0:\bar{k}-1}, \zeta_{0:\bar{k}-1}) \delta x_k + \sum_{k=0}^{\bar{k}-1} \frac{\partial L}{\partial \mathrm{u}_k}(x_{1:\bar{k}}, u_{0:\bar{k}-1}, \zeta_{0:\bar{k}-1}) \delta u_k = 0 \tag{4.17}$$

が任意の微小変動 $\delta x_k \in \mathbb{R}^{n_x}$, $\delta u_k \in \mathbb{R}^{n_u}$ に対して成り立つ

随伴法 (3)：最適性の必要条件

Theorem (最適性の必要条件)

最小化解 $x_{1:\bar{k}}, u_{0:\bar{k}-1}$ に対して，ある $\zeta_{0:\bar{k}-1}$ が存在し，以下が成り立つ．

$$\zeta_{k-1}^\top = \zeta_k^\top \frac{\partial f_k}{\partial \mathrm{x}}(x_k, u_k) + \frac{\partial \ell}{\partial \mathrm{x}}(k, x_k, u_k),\ k = 1, 2, \ldots, \bar{k} - 1 \tag{4.18}$$

$$\zeta_{\bar{k}-1}^\top = \frac{\partial \ell_f}{\partial \mathrm{x}}(x_{\bar{k}}), \tag{4.19}$$

$$\frac{\partial \ell}{\partial \mathrm{u}}(k, x_k, u_k) + \zeta_k^\top \frac{\partial f_k}{\partial \mathrm{u}}(x_k, u_k) = 0,\ k \in \lfloor \bar{k} \rfloor \tag{4.20}$$

- 一般に $\zeta_:$ を随伴変数 (adjoint variable) もしくは共変数 (co-state) とよび，逆時間で発展する．
- (4.20)を u_k について解くことができれば，これを用いてダイナミクスの方程式および (4.18) から u_k を消去して，$x_:, \zeta_:$ に関する二点境界値問題 ($x_0 = \mathrm{x}$ および (4.19)) に帰着する．
- 確率システムの場合は，δx_k, δu_k などが確率変数，等号は期待値の意味となる

線形最適制御

連続状態空間でのベルマン方程式

ベルマン方程式（再掲）

$$V(\bar{k}, \mathbf{x}) = l_f(\mathbf{x}), \forall \mathbf{x} \in \mathbb{X},$$
$$V(k, \mathbf{x}) = \inf_{\mathbf{u} \in U_k(\mathbf{x})} (l + \mathcal{A}_x[V])(k, \mathbf{x}, \mathbf{u}),\ \forall (k, \mathbf{x}) \in [0, \bar{k} - 1] \times \mathbb{X}$$

- $\mathbb{X}$ が $\mathbb{R}^{n_x}$ のように連続集合である場合，ベルマン方程式は無限個の等式条件を持つため近似的にも解くことが困難（次元の呪い）

問題 4.2.1 – 有限時間 LQR（Linear Quadratic Regulator）

線形確率システムに対する二次コスト最小化

$$x_{k+1} = A_k x_k + B_k u_k + v_k,\ \mathbb{E}[v_k] = 0, \mathrm{Var}[v_k] = N_k, \tag{4.21}$$
$$\ell(\mathbf{x}, \mathbf{u}) = \mathrm{x}^\top Q \mathrm{x} + \mathrm{u}^\top R \mathrm{u},\ \ell_f(\mathrm{x}) := \mathrm{x}^\top \Pi_{\bar{k}} \mathrm{x} \tag{4.22}$$

線形二次レギュレータ問題 (2)：LQR 問題の最適解

定理 4.2.2 – 有限時間 LQR 問題の最適解

$$\tilde{R}_k := R_k + B_k^\top \Pi_k B_k,\ \tilde{S}_k := A_k^\top \Pi_k B_k,\ \tilde{Q}_k := Q_k + A_k^\top \Pi_k A_k$$
$$\Pi_{k-1} := \tilde{Q}_k - \tilde{S}_k \tilde{R}_k^{-1} \tilde{S}_k^\top,\ K_k := \tilde{R}_k^{-1} \tilde{S}_k^\top$$

により定まる Π_k は半正定行列であり，つぎの x_0 に依存しない線形制御則が最適解

$$u_k := -K_k x_k \tag{4.23}$$

- $\gamma_{\bar{k}} := 0,\ \gamma_k := \sum_{i=k}^{\bar{k}-1} \mathrm{Trace}(N_i \Pi_{i+1})$ を用いて定義される $V(k, \mathrm{x}) := \mathrm{x}^\top \Pi_k \mathrm{x} + \gamma_k$ が状態価値関数
- **確実性等価原理**（certainty-equivalent principle）：最適制御則が雑音 v_k に依存しない