



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

October 10, 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

動的計画法

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 $f : \mathbb{X} \times \mathbb{U} \times \mathbb{V} \rightarrow \mathbb{X}$
- 初期状態 $x_0 \in \text{rv}(\mathbb{X})$
- \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

$$x_{k+1} = f(x_k, u_k, v_k), k \in \mathbb{Z}_+ \quad (4.2)$$

により定まる確率システム (stochastic system) を扱う

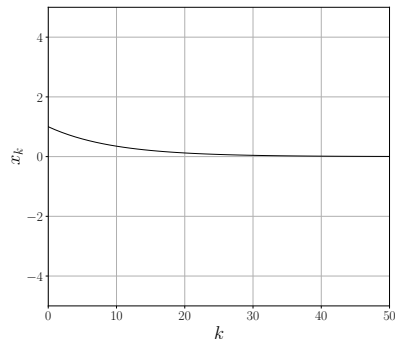
因果性とフィードバック制御 (2) 確定システム

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

$$x(k+1) = ax(k) + u(k), \quad a > 1, \quad x(0) = 1 \quad (4.3)$$

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



確定システム ($a = 1.1$)

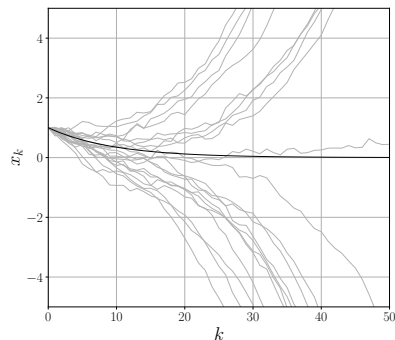
因果性とフィードバック制御 (3) 確率システム

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

$$\begin{aligned} x_{k+1} &= ax_k + u_k + 0.1v_k, \quad a = 1.1, \quad x_0 = 1 \\ v_k &\sim \mathcal{N}(0, 1), \text{ i.i.d.} \end{aligned} \quad (4.4)$$

- $u_k = -Kx_k$ により $\text{Var}[x_k]$ は発散しない
- 任意の確定的な入力に対して $\text{Var}[x_k] \rightarrow \infty$
- 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



確率システム ($a = 1.1$)

因果性とフィードバック制御 (4) 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_k, y_k は速やかに 0 に収束

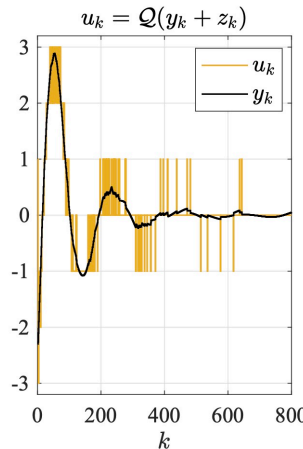
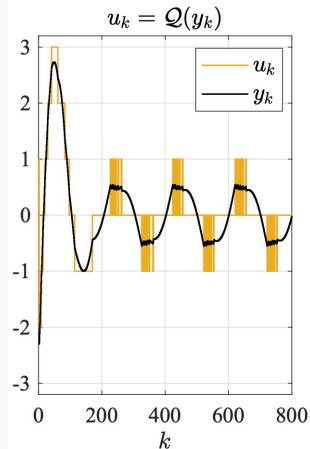
- 最も近い整数を出力する

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

- 確率的状态フィードバック

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

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



最適制御問題 (1) 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

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

- 任意の因果的な入力 u に対して,

$$(x_{0:k}, u_{0:k}) \perp\!\!\!\perp v_k \quad (4.9)$$

- 遷移確率密度関数 $\Psi_k(x_{k+1}|x_k, u_k) := \varphi(x_{k+1}|x_k, u_k)$ は, 関数 f と v_k の分布で定まる

$$\varphi(x_{k+1}|x_{0:k}, u_{0:k}) = \Psi_k(x_{k+1}|x_k, u_k), \forall k, x_{0:k}, u_{0:k} \quad (4.10)$$

最適制御問題 (2) 遷移確率密度関数

例

$$x_{k+1} = f'(x_k, u_k) + g(x_k, u_k)v_k, \quad v_k \sim \mathcal{N}(0, I), \text{ i.i.d.} \quad (4.11)$$

の遷移確率密度関数は,

$$\Psi_k(x'|x, u) = \mathcal{N}(x'|f'(x, u), g(x, u)g(x, u)^\top) \quad (4.12)$$

- 時間発展

$$\varphi(x_{k+1}) = \int_{\mathbb{X}} \int_{\mathbb{U}} \Psi_k(x_{k+1}|x_k, u_k) \varphi(x_k, u_k) du_k dx_k \quad (4.13)$$

- 状態入力系列

$$\varphi(x_{0:k}, u_{0:k-1}) = \varphi(x_0) \prod_{i=0}^{k-1} \left\{ \Psi_i(x_{i+1}|x_i, u_i) \varphi(u_i|x_{0:i}, u_{0:i-1}) \right\} \quad (4.14)$$

- 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

最適制御問題 (3) 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 $x_0 \in \text{rv}(\mathbb{X})$ をもつ確率システム

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

許容入力集合 $U(x) \subset \mathbb{U}$, $k \in [\bar{k}]$, $x \in \mathbb{X}$ に対して,

$$u_k \in U(x_k), \forall k \in [\bar{k}] \quad (4.20)$$

を満たし, コスト関数

$$J(u_\cdot) := \mathbb{E}^u \left[\sum_{k=0}^{\bar{k}-1} \ell(x_k, u_k) + \ell_f(x_{\bar{k}}) \right] \quad (4.21)$$

を最小化する因果的な制御入力 u_\cdot を求めよ.

最適制御問題 (4) 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 $x_0 \in \text{rv}(\mathbb{X})$ をもつ確率システム

任意の因果的な入力 u_{\cdot}^c に対して、対応する状態過程を x_{\cdot}^c とおき、

$$\pi_k^s(u|x) := \varphi_{u_k^c}(u|x_k^c = x), \quad \forall u \in \mathbb{U} \quad (4.23)$$

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

- 各時刻での分布 $\{\varphi_{(x_k, u_k)}\}_{k=0,1,\dots,\bar{k}}$ 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

最適制御問題 (5) 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_k 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

$$\varphi(u_k | x_k) = \pi_k(u_k | x_k) \quad (4.15)$$

で与えられるとき、遷移確率密度関数は

$$\varphi(x_{k+1} | x_k) = \int_{\mathbb{U}} \Psi_k(x_{k+1} | x_k, u_k) \pi_k(u_k | x_k) du_k \quad (4.16)$$

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

$$u_k := \pi_k(x_k) \quad (4.17)$$

で与えられるとき、遷移確率密度関数は

$$\varphi(x_{k+1} | x_k) = \Psi_k(x_{k+1} | x_k, \pi_k(x_k)) \quad (4.18)$$

ベルマン方程式 (1) 状態価値関数

$x_k = x$ の時に達成できる最適累積コスト

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)

$$V^*(k, x) := \inf_{u_k: \in \mathcal{U}} \underbrace{\mathbb{E}_{x_k=x}^u \left[\sum_{i=k}^{\bar{k}} \ell(i, x_i, u_i) \right]}_{=: V^u(k, x)} \quad (4.19)$$

補題 4.1.8 – 状態価値関数と初期状態

$$\inf_{u: \in \mathcal{U}} J(u:.) = \mathbb{E}[V^*(0, x_0)] \quad (4.26)$$

- 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 (2) 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

$$V^*(k, x) = \inf_{u \in U(x)} \{ \ell(x, u) + V^*(k+1, f(x, u)) \} \quad (4.25)$$

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

$$V(\bar{k}, x) = \ell_f(x), \quad \forall x \in \mathbb{X}, \quad (4.28a)$$

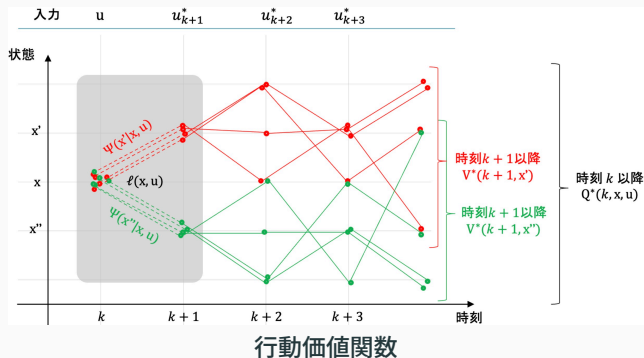
$$V(k, x) = \inf_{u \in U_k(x)} (\ell + \mathcal{A}_x[V])(k, x, u), \quad \forall (k, x) \in [\bar{k}] \times \mathbb{X} \quad (4.28b)$$

$$\mathcal{A}_x[V](k, x, u) := \mathbb{E}[V(k+1, x_{k+1}) | x_k = x, u_k = u] = \int_{\mathbb{X}} V(k+1, x') \Psi_k(x' | x, u) dx' \quad (4.27)$$

ベルマン方程式 (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

$$Q^*(k, x, u) := \inf_{u_{k+1}:} \left\{ \ell(k, x, u) + \mathbb{E}_{x_k=x, u_k=u}^{u_{k+1}:} \left[\sum_{i=k+1}^{\bar{k}} \ell(i, x_i, u_i) \right] \right\} \quad (4.30)$$



ベルマン方程式 (4) 状態価値関数の上界

定理 4.1.11 – 状態価値関数の上界

$$\hat{V}(\bar{k}, x) \geq \ell_f(x), \forall x \in \mathbb{X}, \quad (4.33a)$$

$$\hat{V}(k, x) \geq \inf_{u \in U_k(x)} (\ell + \mathcal{A}_x[\hat{V}])(k, x, u), \forall k \in [\bar{k}], \forall x \in \mathbb{X} \quad (4.33b)$$

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

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

$$\hat{V}(k, x) + \epsilon \geq (\ell + \mathcal{A}_x[\hat{V}])(k, x, \pi_k(x)), \pi_k(x) \in U(x), \forall k \in [\bar{k}], \forall x \in \mathbb{X}, \quad (4.34)$$

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

$$J(\pi) := \mathbb{E}^\pi \left[\sum_{k=0}^{\bar{k}} \ell(k, x_k, u_k) \right] \leq \mathbb{E} \left[\hat{V}(0, x_0) \right] + \epsilon \bar{k} \quad (4.35)$$

ベルマン方程式 (5) 最適制御則

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 (4.33) の等式を満たすことから， $\epsilon = 0$ とす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，

$$\pi_k^*(x) := \arg \min_{u \in U_k(x)} (\ell + \mathcal{A}_x[V^*])(k, x, u) (= \arg \min_{u \in U_k(x)} Q^*(k, x, u)) \quad (4.38)$$

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

ベルマン方程式 (6) 状態価値関数の下界

定理 4.1.13 – 状態価値関数の下界

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

$$\check{V}(\bar{k}, x) \leq \ell_f(x), \quad \forall x \in \mathbb{X}, \quad (4.39a)$$

$$\check{V}(k, x) \leq (\ell + \mathcal{A}_x[\check{V}])(k, x, u), \quad \forall k \in [\bar{k}], x \in \mathbb{X}, u \in U_k(x) \quad (4.39b)$$

を満たすならば, $\check{V} \leq V^*$ である.

随伴法 (1) 確定離散時間システムに対する最適制御問題

確定離散時間システム

$$x_{k+1} = f_k(x_k, u_k), k \in [0, \bar{k} - 1] \quad (4.40)$$

におけるコスト関数

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

を最小化する制御入力を設計する最適制御問題

- 等式拘束条件(4.40)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.
- ラグランジュの未定乗数法の適用

随伴法 (2) 未定乗数法

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

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

$$L(x_{1:\bar{k}}, u_{0:\bar{k}-1}, \zeta_{0:\bar{k}-1}) := \ell_f(x_{\bar{k}}) + \sum_{k=0}^{\bar{k}-1} \{ \ell(k, x_k, u_k) - \zeta_k^\top (x_{k+1} - f_k(x_k, u_k)) \} \quad (4.41)$$

L の停留条件

$$\sum_{k=1}^{\bar{k}} \frac{\partial L}{\partial x_k}(x_{1:\bar{k}}, u_{0:\bar{k}-1}, \zeta_{0:\bar{k}-1}) \delta x_k + \sum_{k=0}^{\bar{k}-1} \frac{\partial L}{\partial u_k}(x_{1:\bar{k}}, u_{0:\bar{k}-1}, \zeta_{0:\bar{k}-1}) \delta u_k = 0 \quad (4.42)$$

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.

$$\zeta_{k-1}^\top = \zeta_k^\top \frac{\partial f_k}{\partial x}(x_k, u_k) + \frac{\partial \ell}{\partial x}(k, x_k, u_k), \quad k = 1, 2, \dots, \bar{k} - 1 \quad (4.43)$$

$$\zeta_{\bar{k}-1}^\top = \frac{\partial \ell_f}{\partial x}(x_{\bar{k}}), \quad (4.44)$$

$$\frac{\partial \ell}{\partial u}(k, x_k, u_k) + \zeta_k^\top \frac{\partial f_k}{\partial u}(x_k, u_k) = 0, \quad k \in [\bar{k}] \quad (4.45)$$

- ζ 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
- (4.45) を u_k について解ければ, $x_{k+1} = f_k(x_k, u_k)$ および (4.43) から u_k を消去して, x_{\cdot}, ζ_{\cdot} に関する二点境界値問題 ($x_0 = x$ および (4.44)) に帰着
- 確率システムの場合は, $\delta x_k, \delta u_k$ などが確率変数, 等号は期待値の意味となる

リスク鋭敏型制御 (1) 定式化

問題 4.1.14 – リスク鋭敏型制御 (risk sensitive control)

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

$$J(u_\cdot) := \mathbb{E}^u \left[\exp \left\{ \tau \left(\sum_{k=0}^{\bar{k}-1} \ell(k, x_k, u_k) + \ell_f(x_{\bar{k}}) \right) \right\} \right] \quad (4.46)$$

を最小化する因果的な制御入力 u_\cdot を求めよ.

- $\tau > 0$ のとき ℓ の実現値の増加に対して評価関数が指数的に増大する**リスク回避型** (risk aversal), $\tau < 0$ のとき**リスク選好型** (risk seeking)
- $(1/\tau) \log \mathbb{E}[\exp \tau \Phi] = \mathbb{E}[\Phi] + \tau \text{Var}[\Phi] + O(\tau^2)$

リスク鋭敏型制御 (2) ベルマン方程式

定理 4.1.15 – リスク鋭敏型制御問題のベルマン方程式

状態価値関数 V^* は，ベルマン方程式

$$V(\bar{k}, x) = \exp(\tau \ell_f(x)), \quad \forall x \in \mathbb{X}, \quad (4.47a)$$

$$V(k, x) = \inf_{u \in U_k(x)} \{ \exp(\tau \ell(k, x, u)) \mathcal{A}_x[V](k, x, u) \}, \quad \forall k \in [\bar{k}], \quad x \in \mathbb{X} \quad (4.47b)$$

の一意解である．

- 後出の LQR と同様に，線形システムかつ二次コストの場合は，行列計算により状態価値関数を構成できる

線形最適制御

線形二次レギュレータ問題 (1) 定式化

ベルマン方程式 (再掲)

$$V(\bar{k}, x) = l_f(x), \forall x \in \mathbb{X}, \quad (4.28a)$$

$$V(k, x) = \inf_{u \in U_k(x)} (l + \mathcal{A}_x[V])(k, x, u), \quad \forall (k, x) \in [0, \bar{k} - 1] \times \mathbb{X} \quad (4.28b)$$

- \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

$$x_{k+1} = A_k x_k + B_k u_k + v_k, \quad \mathbb{E}[v_k] = 0, \text{Var}[v_k] = N_k, \quad (4.48)$$

$$\ell(x, u) = x^\top Q_k x + u^\top R_k u, \quad \ell_f(x) := x^\top \Pi_{\bar{k}} x, \quad Q_k, \Pi_{\bar{k}} \succeq O, R_k \succ O \quad (4.51)$$

線形二次レギュレータ問題 (2) 最適解

定理 4.2.2 – 有限時間 LQR 問題の最適解

$$\tilde{R}_k := R_k + B_k^\top \Pi_k B_k, \tilde{S}_k := A_k^\top \Pi_k B_k, \tilde{Q}_k := Q_k + A_k^\top \Pi_k A_k \quad (4.52)$$

$$\Pi_{k-1} := \tilde{Q}_k - \tilde{S}_k \tilde{R}_k^{-1} \tilde{S}_k^\top, K_k := \tilde{R}_k^{-1} \tilde{S}_k^\top \quad (4.53)$$

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

$$u_k := -K_k x_k \quad (4.56)$$

- $\gamma_{\bar{k}} := 0, \gamma_k := \sum_{i=k}^{\bar{k}-1} \text{Trace}(N_i \Pi_{i+1})$ のもと $V(k, x) := x^\top \Pi_k x + \gamma_k$ 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

第4章

1. 動的計画法

因果性とフィードバック制御

最適制御問題

ベルマン方程式

随伴法

リスク鋭敏型制御

2. 線形最適制御

線形二次レギュレータ問題