

main : 2024/7/13(11:11)

システム制御と統計的学習をつなぐ
確率システム理論

加嶋 健司

2024年7月13日

まえがき

目的

本書はシステム制御に関する理論や応用に携わる方、もしくは分野外の方でシステム制御を勉強してみたい方を対象に、確率的な不確実性を有する動的システムの数学的な取り扱い方を紹介することを目的としています。特に、

1. 制御理論と統計的学習理論（もしくは機械学習・人工知能）の橋渡しとなる知識を整理することと、
2. 確率システム理論の数学的な多様さと美しさを紹介すること、

を念頭におきました。1.に関しては、なんといっても近年、データサイエンスが改めて大きな注目を集めていることが背景にあります。古くから自動制御と人工知能は関連の深い分野ですが、「強化学習はモデルを用いない制御系設計である」、「システム同定は教師あり学習の一種である」といった表層的な合意はただちに得られるものの、そこからなかなか議論が深まらないこともあります。この大きな機会損失は、分野間の文化的な違いを語る以前に、数学的な概念の共有が不十分であることが一因であり、確率システムに関する著者の拙い知識を共有することでも、この距離をある程度埋めることができると感じました。機械学習手法はブラックボックスであるため信頼性が足りないという懸念に対する、地に足のついた対処になることも期待しています。2.に関しては、数学的に厳密な多くの良書が存在する一方で、確率システムを特別に研究テーマに選んだ方以外には敷居が高い場合が多いという声もよくきます。また、多様性については、個別のテーマが重厚かつ視点が大きく異なるため、結

果を積み上げる形で確率システム理論の全体像を概観することは容易ではありません。そこで、本書ではこうした各論を、最新の結果まで含めて、広く紹介する形式としました。ただし、数学的な厳密さはシステム制御理論のもつ大きな強みですので、本書のみで議論が閉じるよう配慮しています。両分野の交流において、例えば機械学習の様々な概念や証明技法が、システム制御の課題や結果と紐づけて理解していただければ幸甚です。

内 容

まず本書を通じて、離散時間システムをあつかうことにしました。これは、上述の敷居の高さの原因を、連続時間ダイナミクス（確率微分方程式）の扱いに起因するものと、確率システム特有の理論展開に起因するものに明確に分離したかったためです。これにより、データの取得や勾配法の更新など元来離散時間的に実装される現象の記述とは相性がよくなる一方で、結果の美しさが損なわれたり、統計力学的な視点など見通しが悪くなるという欠点を受け入れました。また、本書は解釈に多くの紙幅を費やす離乳食的な書き方はしておらず、必要であれば数学的に厳密な議論する素地が身につくことを重視しました（必要であれば一から理論構築できるよう、測度論的確率論にも触っています）。導出に用いる論法も、標準的なシステム制御理論とはかなり異なる部分もあるため、併せて楽しんで頂ければと思います。

主な読者として、大学3回生以上の数理工学に関心のある学生や、学部や高専で制御理論を履修し、それを土台として機械学習をはじめとする隣接分野に視野を広げたい方、長年更新されていない現代制御の講義内容を検討する材料をお探しの先生方などを想定しました。より対象を具体化したい箇所では、著者の指導学生が「卒業研究でシステム制御と機械学習の融合研究に取り組む」ことを希望している場合に、できるだけ広い視野と賞味期限の長い知識を与える読み物となることを目指しました。

4

○○○○○○○○○○○○○○○○○○○○○○○○○○

大学 ●●

謝 辞

山本先生、藤岡先生、井村先生、大塚先生、太田先生、足立先生
手伝ってくれた人

main : 2024/7/13(11:11)

目 次

第I部 基礎編

xiii

1. はじめに (6)

1.1 本書の構成	1
1.2 準備	2

2. 確率と統計 (19)

2.1 確率分布モデル	8
2.2 統計量	11
2.2.1 代表的な統計量	11
2.2.2 多変量正規分布	15
2.3 統計的推論	16
2.3.1 ベイズ統計	16
2.3.2 点推定	20

3. 確率システム (17)

3.1 確率過程	27
3.1.1 マルコフ過程	27
3.1.2 分布の時間発展	31

目 次 vii

3.1.3 リアノフ安定論	33
3.2 線形確率システム解析	36
3.2.1 時 間 領 域	36
3.2.2 周 波 数 領 域	40

4. 確率制御 1 (18)

4.1 動的計画法	44
4.1.1 因果性とフィードバック制御	44
4.1.2 ベルマン方程式	51
4.1.3 【オプション】リスク鋭敏型制御	54
4.2 最適制御則	55
4.2.1 ベルマン方程式の不等式評価	55
4.2.2 線形二次レギュレータ	58
4.2.3 【オプション】最適性原理	60

5. 状態推定 (17)

5.1 状態推定	62
5.1.1 統計的推論としての定式化	62
5.1.2 線形システムの可観測性と最尤初期状態推定	63
5.2 逐次推定	65
5.2.1 線形システムの逐次予測	66
5.2.2 線形二次ガウシアン制御	68
5.2.3 カルマンフィルタ	70
5.3 軌道の MAP 推定と最適制御の双対性	73

6. 最適制御 2 (18)

6.1 動的計画法	79
6.1.1 定式化	79
6.1.2 ベルマン方程式	81
6.1.3 儲値反復法	85
6.2 線形制御	87
6.2.1 安定性	87
6.2.2 無限区間線形二次レギュレータ	88
6.2.3 【オプション】長時間平均コスト	93

第 II 部 発展編	97
------------	----

7. 統計的学習(18)

7.1 関数回帰	99
7.1.1 統計推論としての定式化	99
7.1.2 最小二乗法と正則化	102
7.2 特徴量抽出	108
7.2.1 主成分分析	108
7.2.2 次元削減	110
7.3 標本の生成と利用	112
7.3.1 標本生成	112
7.3.2 モンテカルロ法	113
7.3.3 生成モデル	115

8. システム同定 (19)

8.1 最小二乗法	117
8.1.1 推定器としての最小二乗法	117
8.1.2 システム同定	121
8.2 スペクトル法	125
8.2.1 正準相関解析と確率実現	125
8.2.2 モデル低次元化と動的モード分解	130

9. 強化学習 (18)

9.1 最適制御と強化学習	136
9.2 モデルベースト設計	138
9.2.1 方策反復法	138
9.2.2 方策勾配法	143
9.3 モデルフリー設計	145
9.3.1 行動価値関数の推定	145
9.3.2 アルゴリズム例	149
9.3.3 部分観測問題	152

10. マルコフ連鎖と平衡状態 (19)

10.1 有限状態確率システム	154
10.1.1 ベルマン方程式と線形計画法	154
10.1.2 線形可解マルコフ決定過程	157
10.2 漸近拳動	162

x 目 次

10.2.1 ペロン・フロベニウスの定理	162
10.2.2 定常状態と平衡状態	165

11. 数理最適化 (16)

11.1 半正定値計画問題	171
11.1.1 ベルマン方程式の不等式評価	171
11.1.2 状態制約付き線形制御	176
11.2 確率勾配法	179
11.2.1 誤差逆伝播	179
11.2.2 確率近似法	182

12. ガウス過程回帰 (16)

12.1 カーネル法	187
12.1.1 半正定値カーネル	187
12.1.2 ガウス過程回帰	190
12.1.3 【オプション】再生核ヒルベルト空間	194
12.2 ガウス過程回帰状態空間モデル	199
12.2.1 モーメントマッチング	200
12.2.2 ベイズ最適化	202

13. おわりに

付録

目 次 *xi*

A	測度論的確率論	205
A.1	確率と確率変数	205
A.2	条件付き期待値・確率	212
A.3	確率過程	214
A.4	定常確率過程	216
B	定理の証明	220
B.1	フィッシャー情報行列	220
B.2	生成モデルの敵対的学习	222
B.3	方策勾配定理	224
B.4	マルコフ連鎖の漸近特性	226
B.5	確率勾配法	227
引用・参考文献		230

main : 2024/7/13(11:11)

main : 2024/7/13(11:11)

第I部

基礎編

main : 2024/7/13(11:11)

- \mathbb{R} は実数の集合
- \mathbb{Z} は整数, \mathbb{Z}_+ は非負整数の集合
- $X \times Y$ は X と Y の直積集合 $\{(x, y) : x \in X, y \in Y\}$
- $\llbracket l, k \rrbracket := \{l, l+1, \dots, k\}$, $\llbracket k \rrbracket := \llbracket 1, k \rrbracket$, $\llbracket k \rrbracket := \llbracket 0, k-1 \rrbracket$
- $A \succ B$ ($A \succeq B$) は $A - B$ は (半) 正定行列
- $\text{rv}(X)$ は X に実現値をとる確率変数の集合
- \mathbb{P} は確率, \mathbb{E} は期待値, φ_x は確率変数 x の確率密度関数
- \mathbb{X} は状態変数, \mathbb{U} は入力, \mathbb{Y} は出力, \mathbb{V} はシステム雑音, \mathbb{W} は観測雑音が取りうる値の集合
- \mathcal{U}_a は因果性かつ状態依存入力制約を満たす入力過程の集合
- $\mathcal{N}(\cdot | \cdot, \cdot)$ は正規分布, $\text{Lap}(\cdot | \cdot, \cdot)$ はラプラス分布, $\text{Uni}(\cdot)$ は一様分布
- $\delta_x(\cdot)$ はディラックのデルタ関数, $\mathbb{1}_B(\cdot)$ は指示関数
- 関数 $f: X \rightarrow \mathbb{R}$ に対して, $\|f\|_\infty := \sup_{x \in X} |f(x)|$
- $\langle \cdot, \cdot \rangle_X$ は内積空間 X における内積
- C_k は可制御性行列, $G_k^r := C_k C_k^\top$ は可到達性グラミアン
- O_k は可観測性行列, $G_k^o := O_k^\top O_k$ は可観測性グラミアン

主な確率変数と関数の役割 (括弧内は定数)

- p (p) は線形パラメトリック表現のパラメータ
- q (q) は特徴ベクトル
- x (x) は状態変数, 回帰問題の代入点
- y (y) は出力, 観測値
- u (u) は入力
- v (v) はシステム雑音
- w (w) は観測雑音
- f (f) は状態遷移関数, 回帰関数
- g は
- h は出力関数
- k は時間変数 (\bar{k} は終端時刻), カーネル関数

main : 2024/7/13(11:11)

xvi _____

- ϕ は特徴写像

1|| はじめに (6)

1.1 本書の構成

システム制御理論の標準的な用語にもとづいて、各章の内容とそれらの関係を図 1.1 にまとめた。まず、本書で用いる確率統計の基礎知識を 2 章で整理したのち、3 章から 6 章までで一般的な現代制御論の内容 [1,2] に対応する確率システムの場合の結果が概観できる。具体的には、3 章では時刻とともに変化する現象（ダイナミクス）の基本性質、4 章では最適制御、5 章では状態推定、6 章では安定性など無限時間区間の挙動、をそれぞれ取り扱う。これらの章での証明技法などは標準的な現代制御論と類似しているため、適宜比較しつつ確率システム特有の議論に慣れて欲しい。また、統計的推論の概念をシステム制御の視点から理解することも有用であると思う。

続く三つの章は、データの取り扱いと関連が深い。7 章では、一般に機械学習と呼ばれる統計的学習の理論と手法の基礎を概説する。また、この章の内容を用いて、システム同定、つまり動的システムのモデル構築に関する結果を 8 章にて導出する。9 章では、非モデルベーストの制御系設計手法である強化学習の考え方を紹介する。基礎となるモデルベーストの理論（6 章）が前提であり、一部でシステム同定の考え方も援用する。

残る章はほぼ独立しているため、興味のあるトピックのみ読むことができる。

- 10 章では、通常のシステム制御理論では扱わないが、機械学習分野では

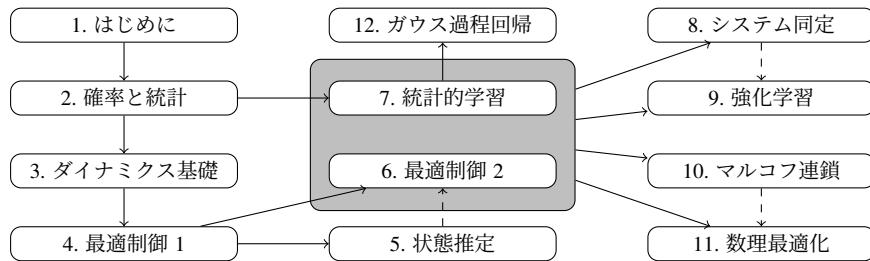
2 1. はじめに (6)

図 1.1: 各章の関係

標準的な有限個の値しか取らない確率システムに限定した話題を紹介する。その後半では、連続状態の厳密な議論が数学的に難しい定常状態に関する結果もまとめる。

- 近年の制御理論分野では、設計問題を凸最適化問題に帰着させることを目標に据える研究が多いが、11章の前半では確率システムに対する同様の結果を与える。後半では対照的に、機械学習分野では重要な役割を果たす確率勾配法の考え方を紹介する。
- 12章では、単純な最小二乗法の拡張とも言えるガウス過程回帰を概説し、動的システムの表現に用いる際の留意点をまとめた。

最後に、本書の内容の多くは、連続時間システムの場合でも同様の結果が成立つ。13章において、これらを展望とともに述べる。

ところで、理工学の様々な場面において不確実性の適切な評価と対処は重要な課題であるが、不確実性の定量化は容易ではなく、過度に解釈に振り回されることも多い。そこで、1.2節において本書で用いる確率に関する前提を整理する。

1.2 準 備

「このサイコロは集合 $\{1, 2, 3, 4, 5, 6\}$ に実現値をもつ確率変数である」という慣れ親しんだ表現と同様に、集合 $X \subset \mathbb{R}^d$ ($d \in \mathbb{N}$) に対して「 x は X に実現値

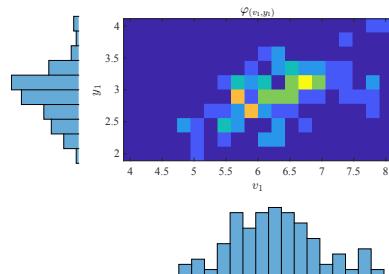


図 1.2: 結合密度関数と周辺密度関数

をもつ確率変数 (random variable) である」ことを、英語名の頭文字をとって

$$x \in \text{rv}(X) \quad (1.1)$$

と表記する。本書では、確率変数の厳密な定義は割愛し^{†1}、つぎの性質を前提として認めることにする。

前提 (確率変数・確率密度関数)。任意の $x \in \text{rv}(X)$ に対して、規格化条件^{†2}

$$\int_X \varphi_x(x) dx = 1 \quad (1.2)$$

を満たす確率密度関数 (probability density function) と呼ばれる非負値関数 $\varphi_x : X \rightarrow \mathbb{R}_+$ が唯一つ存在し、この関係を $x \sim \varphi_x$ と表記する。さらに、任意の $x \in \text{rv}(X)$, $y \in \text{rv}(Y)$ に対して $(x, y) \in \text{rv}(X \times Y)$ であり、

$$\varphi_x(x) = \int_Y \varphi_{(x,y)}(x, y) dy, \quad \varphi_y(y) = \int_X \varphi_{(x,y)}(x, y) dx \quad (1.3)$$

などが成り立つ。 \diamond

例 1.2.1. ある風力発電所での時刻 k における風速と発電量を確率変数 $v_k, y_k \in \text{rv}(\mathbb{R})$ によりそれぞれ表現すると、 $(v_1, y_1) \in \text{rv}(\mathbb{R}^2)$, $(y_1, y_2, \dots, y_{10}) \in \text{rv}(\mathbb{R}^{10})$ である。図 1.2 のように (v_1, y_1) の確率密度関数 $\varphi_{(v_1, y_1)}$ が各長方形領域ごとに

^{†1} 密度関数の存在を仮定しない測度論にもとづく定義を付録 Aにおいて紹介する。

^{†2} 以降の積分は、通常の \mathbb{R}^d におけるリーマン積分であると思ってよい。

4 1. はじめに (6)

定数関数である場合を考えると、馴染みのあるサイコロを二投する場合との類似性がわかるだろう。ここで、周辺（下と左）に示す各軸方向の和をとった関数が φ_{v_1} , φ_{y_1} である。また、 $\varphi_{(v_1,y_1)}$ が大きい値をとる領域が右上がりに分布していることは、風が強いほど発電量が多いことを示唆している。 ◇

(1.3) のような積分操作により一部を消去した確率変数の密度関数を抽出することを周辺化 (marginalization), 周辺化の前／後の確率密度関数を結合／周辺確率密度関数 (joint / marginal) と呼ぶ。「考察の対象とするすべての個別の確率変数 $\{x_1, x_2, \dots\}$ の結合確率密度関数 $\varphi_{(x_1, x_2, \dots)}$ が唯一つ準備されており、その周辺化により任意の確率変数の密度関数が定まる」というイメージをもつことができれば、ひとまず先にすすんでほしい。

定義 1.2.2 (確率・期待値). $x \in \text{rv}(X)$, $B \subset X$ に対して

$$\mathbb{P}_x(B) := \int_B \varphi_x(x) dx \quad (1.4)$$

を x の実現値が集合 $B \subset X$ に含まれる確率 (probability), 関数 \mathbb{P}_x を x の分布 (distribution),

$$\mathbb{E}[x] := \int_X x \varphi_x(x) dx \quad (1.5)$$

を x の期待値 (expectation) と呼ぶ。 ◇

注意 1.2.3 (書体の使い分け). 本書を通じて,

$$x \in \text{rv}(X) \text{ の実現値 } x \in X \quad (1.6)$$

といった書体の使い分けを採用する。また、実現値が満たす条件は、たとえば $\mathbb{P}_x(\{x : |x| > 1\})$ を $\mathbb{P}(|x| > 1)$ のように略記する。特に、 $\mathbb{P}(x = x)(= \mathbb{P}_x(\{x\}))$ は「 x が特定の実現値 x をとる」確率と解釈してよい。 ◇

離散値をあつかう場合と比較すると、連続値をあつかうため和が積分に置き換わっているが、いずれも「 $\varphi_x(x)$ は $x \in X$ の実現しやすさ, $\mathbb{P}_x(B)$ は B における

φ_x の累積値, $\mathbb{E}[x]$ は実現値の重み付き平均」などと容易に受け入れられる関係式であろう。それでも、確率 $\mathbb{P}(x = x)$ と確率密度関数 $\varphi_x(x)$ などは混同しやすいかもしない。例えば, $\mathbb{P}_x(B) \leq 1, \forall B \subset X$ であるが φ_x はいくらでも大きい値をとりうる。あるいは φ_x が滑らかな関数であるとき, $\mathbb{P}(x = x) = 0, \forall x \in X$ であるが、「 $\varphi_x(x) = 0$ ($\varphi_x(x) > 0$) となる x に対しては, $x = x$ は実現可能性がない（ある）」という感覚で概ね正しい。

確率変数 $x, y \in \text{rv}(X)$ の等号 $x = y$ は $\mathbb{P}(x = y) (= \mathbb{P}_{(x,y)}(\{(x,y) : x = y\})) = 1$ の意味で用いるが, 素朴に「 x と y の実現値は常に等しい」と考えればよく, 不等号 $x < y$ などについても同様である。また, $x \in \text{rv}(X)$ および関数 $f : X \rightarrow Y$ に対して, $y := f(x) \in \text{rv}(Y)$ は「 x の実現値が x のとき常に実現値 $f(x)$ をとする確率変数」である。これにより, 確率変数 x からスカラー倍 ax ($a \in \mathbb{R}$) や e^x などの別の確率変数を生成したり, 和 $x + y$ など確率変数の演算をおこなうことが自然に定義される[†].

つぎに確率変数間の依存関係に関する概念を整理する。

定義 1.2.4 (独立・同分布). $x_i \in \text{rv}(X_i), i = 1, 2, \dots, l$ に対して, $x = (x_1, \dots, x_l)$ の表記のもと,

$$\varphi_x(x) = \prod_{i=1}^l \varphi_{x_i}(x_i), \forall x_i \in X_i, i = 1, \dots, l \quad (1.7)$$

が成り立つとき, $\{x_i\}_{1 \leq i \leq l}$ は独立 (independent) であるという。特に, x, y が独立であることを $x \perp\!\!\!\perp y$ と表記する。一方, $x, y \in \text{rv}(X)$ が $\varphi_x = \varphi_y$ を満たすとき, x と y は同分布をもつ (identically distributed) といい, $x \stackrel{d}{=} y$ と表記する。また, $y = f(x)$ を満たす $f : X \rightarrow Y$ が存在するとき,

$$y \in \sigma(x) \quad (1.8)$$

と表記する。 ◁

独立性は確率変数 x_i が影響を及ぼし合わないという状況の自然な表現であり, $x \perp\!\!\!\perp y$ ならば $\mathbb{E}[xy] = \mathbb{E}[x]\mathbb{E}[y]$ であることが容易に確かめられる。なお, 任意

[†] この演算のもと, $\text{rv}(\mathbb{R}^n)$ は内積 $\langle x, y \rangle := \mathbb{E}[x^\top y]$ を備えた内積空間とみなせる。

6 1. はじめに (6)

の組 $1 \leq i < j \leq l$ に対して $x_i \perp\!\!\!\perp x_j$ であっても、全体として $\{x_i\}_{1 \leq i \leq l}$ が独立とは限らない。また、(1.8) の記法は、「状態 x を用いた（確定的）状態フィードバック入力 $u \in \sigma(x)$ 」や「出力 y にもとづく状態推定値 $\hat{x} \in \sigma(y)$ 」など、因果性を記述する際に多用する。もちろん、 $u \in \sigma(\hat{x})$, $\hat{x} \in \sigma(y)$ ならば、 $u \in \sigma(y)$ である。また、 $x = y$ ならば x と y は同分布をもつが、その逆は成り立たない。これは確率密度関数 φ_x , φ_y には (x, y) の結合確率密度関数の情報が含まれていないためであり、たとえば、集合 $B \subset X$ の指示関数 $\mathbb{1}_B : X \rightarrow \{0, 1\}$

$$\mathbb{1}_B(x) := \begin{cases} 1, & x \in B \\ 0, & x \notin B, \end{cases} \quad (1.9)$$

を用いて、 $x \in \text{rv}(\mathbb{R})$ の確率密度関数が $\varphi_x(x) \propto \mathbb{1}_{(-1,1)}(x)$ と与えられるとき[†]、 x と $y := -x$ は同分布をもつが[‡]、 $\mathbb{P}(x = y) = 0$ である。なお、 $x \perp\!\!\!\perp y$ の仮定のもとでは、 φ_x , φ_y から $\varphi_{(x,y)} = \varphi_x \varphi_y$ が定まる。

最後に、ここまでやや形式的に導入してきた確率の概念が、直感的にも納得できる性質を有することを、定理として（証明はせずに）紹介しておく。

定理 1.2.5 (従属確率変数の独立性). 確率変数 x, y, z が $x \perp\!\!\!\perp y$ かつ $z \in \sigma(y)$ を満たすならば、 $x \perp\!\!\!\perp z$ である。 ◇

定理 1.2.6 (Law of the unconscious statistician). 任意の $x \in \text{rv}(X)$ と関数 $f : X \rightarrow Y$ に対して、 $f(x) \in \text{rv}(Y)$ の期待値 $\mathbb{E}[f(x)]$ は

$$\mathbb{E}[f(x)] \left(:= \int_Y y \varphi_{f(x)}(y) dy \right) = \int_X f(x) \varphi_x(x) dx \quad (1.10)$$

を満たす。 ◇

定理 1.2.7 (期待値演算の線形性). 任意の $x_1, x_2 \in \text{rv}(\mathbb{R}^n)$, $a_1, a_2 \in \mathbb{R}$ に対して、

$$\mathbb{E}[a_1 x_1 + a_2 x_2] = a_1 \mathbb{E}[x_1] + a_2 \mathbb{E}[x_2] \quad (1.11)$$

が成り立つ。 ◇

[†] 本書を通じて、重要でない規格化定数を省略するために比例をあらわす記号 \propto を用いる。例えば、 $\varphi_x(x) \propto \mathbb{1}_{(-1,1)}(x)$ は $\varphi_x(x) = 0.5 \cdot \mathbb{1}_{(-1,1)}(x)$ である。

定理 1.2.8 (独立な確率変数の和の確率密度関数). $x, y \in \text{rv}(\mathbb{R}^n)$ が $x \perp\!\!\!\perp y$ を満たすとき, $z := x + y$ は

$$\varphi_z(z) = (\varphi_x * \varphi_y)(z) := \int_{\mathbb{R}^n} \varphi_x(x) \varphi_y(z - x) dx \quad (1.12)$$

を満たす. \diamond

定理 1.2.6において f を $B \subset X$ の指示関数 $\mathbb{1}_B$ とすると, (1.4) は

$$\mathbb{P}_x(B) = \mathbb{E}[\mathbb{1}_B(x)] \quad (1.13)$$

とも書ける. 定理 1.2.7 は定理 1.2.6 からしたがう. 定理 1.2.8 は, 被積分関数が $\varphi_{(x,y)}(x, z-x)$ であり, その x に関する累積値が $\varphi_{x+y}(z)$ と理解すればよい.

なお, $\{0,1\}$ の二値を等確率でとる確率変数をコイン投げで実装するように, 与えられた X -上の確率密度関数に比例する割合^{†1}で $x \in X$ を抽出する手続きを一般に（疑似）乱数生成と呼び, 生成された x を標本（sample）または乱数（random number）と呼ぶ. こうした乱数生成にもとづくシミュレーションは, 確率モデルの可視化やその性質を調べる有力な手法である^{†2}.

本書を読み進めるために想定する確率に関する事前知識は, 以上すべてである.

^{†1} 本書の議論は測度論的確率論にもとづいており, 頻度論的な内容はあつかわない.

^{†2} 亂数生成にもとづくモンテカルロ法に関する結果を 7.3.2 節にて紹介する.

2 確率と統計 (19)

一般的にモデルとデータの関係性について「確率と統計」という見方をすることが多い。そこで本章では、2.1節にて確率モデルを具体的に準備する方法、2.2節において確率モデルから定まる実現値の傾向を定量化する確率論に近い考え方、2.3節において実現値から確率モデルを定める統計学に近い考え方、のそれぞれに関して必要な数学的な概念を、最小限の説明[†]で列挙する。

2.1 確率分布モデル

本節では、 $x \in \text{rv}(\mathbb{R}^n)$ の密度関数 φ_x を具体的に準備する方法について考察する。ここで、確率密度関数は規格化条件 (\mathbb{R}^n での積分値が 1) を満たす非負値関数であれば任意でよいが、特別な理由がない限り、確率分布モデルと呼ばれる特定の関数形が採用される。有界な集合 $X \subset \mathbb{R}^n$ に対して、 $x \in \text{rv}(X)$ の確率密度関数が定数関数であるとき、 x は**一様分布** (uniform distribution) にしたがうといい、 $x \sim \text{Uni}(X)$ とあらわす（規格化条件よりすべての実数が等確率で実現する確率変数は定義できない）。また、 $a \in \mathbb{R}^n$ に対して $\mathbb{P}(x = a) = 1$ であるとき $x \sim \delta_a$ とあらわし、その確率密度関数を**ディラックのデルタ関数** (Dirac's delta) $\delta_a(\cdot)$ と表記する。これは、任意の滑らかな関数 f に対して

$$\mathbb{E}[f(x)] = \int f(x) \delta_a(x) dx = f(a), \quad (2.1)$$

[†] 標準的な内容 [3] であるため、より詳しい説明や具体例が必要な場合も、趣向を凝らした解説記事が容易に見つけられるはずである。

特に、任意の $B \subset \mathbb{R}^n$ に対して

$$\int_B \delta_a(x) dx = \mathbb{1}_B(a) \quad (2.2)$$

を満たす（超）関数と思えばよい。

数ある確率分布モデルのなかでも、正規分布は单峰性や対称性をもち $e^{-c\|x\|^2}$ ($c > 0$) のオーダで減衰する確率密度関数しか表現できないものの、様々な数学的に有益な性質を有するため、一般的に用いられる。

定義 2.1.1 (正規分布). $x \in \text{rv}(\mathbb{R}^n)$ の確率密度関数が $\mu \in \mathbb{R}^n$ および正定行列 $\Sigma \in \mathbb{R}^{n \times n}$ を用いて

$$\mathcal{N}(x|\mu, \Sigma) := \frac{1}{(2\pi)^{n/2} \sqrt{\det \Sigma}} \exp\left(-\frac{1}{2}(x - \mu)^\top \Sigma^{-1}(x - \mu)\right) \quad (2.3)$$

で与えられるとき、 x は平均 μ 、共分散 Σ をもつ多変量の**正規分布** (multivariate normal distribution) にしたがうといい、 $x \sim \mathcal{N}(\mu, \Sigma)$ とあらわす。 \triangleleft

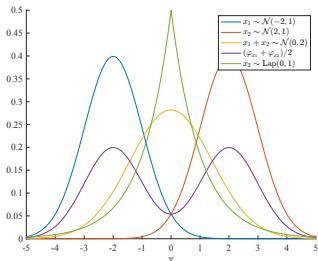
他の例としては、確率密度関数の減衰が遅い（裾が厚い、といわれる）分布、すなわち外れ値 (outlier) と呼ばれる平均値から大きく離れた値をとる確率が非常に大きい場合によく用いられる確率分布モデルにつきのラプラス分布がある。

定義 2.1.2 (ラプラス分布). $\alpha \in \mathbb{R}$, $\beta > 0$ に対して、 $x \in \text{rv}(\mathbb{R})$ の確率密度関数が

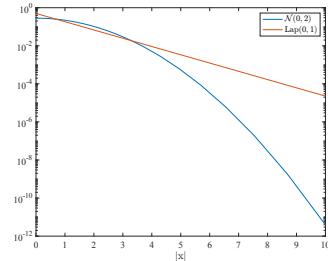
$$\text{Lap}(x|\alpha, \beta) := \frac{1}{2\beta} \exp\left(-\frac{|x - \alpha|}{\beta}\right) \quad (2.4)$$

で与えられるとき、 x は**ラプラス分布** (Laplace distribution) にしたがうといい、 $x \sim \text{Lap}(\alpha, \beta)$ とあらわす。 \triangleleft

(2.4) の対数をとると区分線形関数となることは状態推定などをおこなう際に有益（例 5.3.3 参照）である一方、再生性（定理 2.2.10）をもたないなど、正規分布ほど扱いやすくはない。正規分布を直接的に拡張した**安定分布**は、いわゆるべき乗則 $|x|^{-c}$ ($c > 0$) にしたがうさらに裾の厚い分布を表現でき、正規分布と

10 2. 確率と統計(19)

(a) 正規分布の再生性と混合正規分布



(b) 分布の裾

図 2.1: 正規分布とラプラス分布

類似したいくつかの望ましい性質を有するが、確率密度関数が解析的に書き下せないなど、その理論的に厳密な取扱いには注意を要する [4,5]。分布が多峰性を有する場合は、

$$\varphi_x(x) = \sum_i a_i \mathcal{N}(x|\mu_i, \Sigma_i), \quad a_i \geq 0, \quad \sum_i a_i = 1 \quad (2.5)$$

であらわされる混合ガウス分布も広く用いられる（図 2.1a）。なお、(2.5) は確率密度関数の線形結合であり、独立な正規分布にしたがう確率変数の線形結合（定理 1.2.8 参照）ではない。

定義 2.1.1 の μ, Σ や、定義 2.1.2 の α, β のように分布を特徴づけるパラメータそのものが別の確率変数として与えられる階層的なモデルもしばしば用いられる。たとえば、

$$\varphi_{(z,x,y)}(z, x, y) \propto \mathcal{N}(z|x, y) \cdot \text{Lap}(x|0, 1) \cdot \mathbb{1}_{[1,2]}(y) \quad (2.6)$$

とすれば、「 z は正規分布にしたがうが、その平均 x は裾の厚い分布、分散 y も区間 $[1, 2]$ 上の一様分布から確率的に定まる」と解釈できる。このとき、 x, y を周辺化することで z の密度関数 φ_z は定まるが、正規分布にもラプラス分布にもしたがわない。なお、この例で「 x, y の実現値により z の分布が定まる」といった関係を図 2.2 のようにあらわしたもの **グラフィカルモデル** と呼ぶ。これは $z \in \sigma(x, y)$ とは違うので注意してほしい。

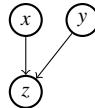


図 2.2: グラフィカルモデル

本書ではごく限られた確率分布モデルと考え方のみを紹介するが、統計的学習分野においては、様々な確率分布モデルが用いられる [6]。確率モデルの選択は得られる結果に大きく影響するため、表現したい現象や利用目的を踏まえて慎重に決定する必要がある[†]。

2.2 統 計 量

風力発電量を見積るために「地点 A, B における風速」という物理量が x_A, x_B という確率変数でモデル化されているとしよう。このとき、どれほどの変動が見込まれるか、地点 B で風が強いとき地点 A でも強いのか、といった質問にどうすれば客観的に答えられるだろうか。本節では、こうした確率変数の実現値の振る舞いを定量的に特徴づけるための代表的な指標とその性質を紹介する。

2.2.1 代表的な統計量

(1) 確率密度関数の形状 確率変数 x に対して、確率密度関数の特徴量、すなわち φ_x で定まる値を x の統計量と呼ぶ。 φ_x の特徴を説明するうえで、その重み付き平均にあたる $\mathbb{E}[x]$ はやはりもっとも基本的な指標であり、この値からの変動の大きさが後出の分散行列である。確率密度関数を最大化する

$$\text{Mode}[x] := \arg \max_{x \in \mathbb{R}^n} \varphi_x(x) \quad (2.7)$$

は確率変数 x のモード（最頻値）と呼ばれる。分布の形状が单峰性・対称性などを有さない限りは、モードと平均は乖離する。

また、 φ_x の（多次元）フーリエ変換はつぎのように定義される。

[†] ??節において、モデルに含まれるパラメータの決定法について考察する。

12 2. 確率と統計(19)

定義 2.2.1 (特性関数). $x \in \text{rv}(\mathbb{R}^n)$ に対して,

$$\Phi_x(\lambda) := \mathbb{E}[\exp(j\lambda^\top x)] = \int_{\mathbb{R}^n} \exp(j\lambda^\top x) \varphi_x(x) dx, \quad \lambda \in \mathbb{R}^n \quad (2.8)$$

を確率変数 x の**特性関数** (characteristic function) と呼ぶ. \triangleleft

フーリエ変換は全単射写像であるため、同分布をもつことは特性関数の一致性とも言い換えられる。規格化条件および $|\exp(j\lambda^\top x)| \leq 1$ より、 $|\Phi_x(\lambda)| \leq 1, \forall \lambda \in \mathbb{R}^n$ 、特に $\Phi_x(0) = 1$ である。畳み込み積を通常の積へと変換することが古典制御でのフーリエ変換（ラプラス変換）の利用方法であったが、定理 1.2.8 の性質を

$$\Phi_{x+y}(\lambda) = \Phi_x(\lambda)\Phi_y(\lambda), \quad \forall \lambda \in \mathbb{R}^n \quad (2.9)$$

として扱うことは、再生性（定理 2.2.10）や中心極限定理の証明など独立な確率変数の和の解析の常套手段である。

つぎのエントロピーは、情報理論や統計力学的な視点において中心的な役割を果たす「乱雑さ」の指標である。

定義 2.2.2 (エントロピー). \mathbb{R}^n -値確率変数 x に対して,

$$H(x) := \int_{\mathbb{R}^n} -\varphi_x(x) \log \varphi_x(x) dx \quad (2.10)$$

を x の**エントロピー** (entropy) と呼ぶ. \triangleleft

直感的には、確率密度関数が特定の箇所でのみ大きな値をとるほどエントロピーは小さくなり（乱雑さが低い）、定数関数に近いほど大きな値をとる。

(2) 確率変数間の関係 ここでは確率変数間の関係を定量的に特徴づける。まず、確率密度関数の近さに関しては、つぎの指標がよく用いられる。

定義 2.2.3 (カルバッック・ライブラ情報量). $x, y \in \text{rv}(\mathbb{R}^n)$ が、 $\varphi_y(x) = 0$ である任意の $x \in \mathbb{R}^n$ に対して $\varphi_x(x) = 0$ を満たすとき

$$x \ll y \quad (2.11)$$

とあらわす。また、(2.11) を満たす x, y のカルバッック・ライブラ情報量 (Kullback-Leibler divergence) または相対エントロピー (relative entropy) を

$$D_{\text{KL}}(x||y) := \int_{\mathbb{R}^n} \varphi_x(x) \log \frac{\varphi_x(x)}{\varphi_y(x)} dx \quad (2.12)$$

により定義^{†1}する。 \triangleleft

カルバッック・ライブラ情報量は、カルバッック・ライブラ距離とも呼ばれるが、距離関数の公理^{†2}のうち対称性と三角不等式は満たさない。ここで、イエンセンの不等式を用いて期待値を評価する論法を紹介する形で、カルバッック・ライブラ情報量の非負値性を証明しておこう。

定理 2.2.4 (イエンセン (Jensen) の不等式). 任意の $x \in \text{rv}(\mathbb{R}^n)$, 関数 $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$ および凸関数 $h: \mathbb{R}^m \rightarrow \mathbb{R}$ に対して, $g(x)$ の期待値が有限ならば,

$$\mathbb{E}[h(g(x))] \geq h(\mathbb{E}[g(x)]) \quad (2.13)$$

が成り立つ。 \triangleleft

証明. $\mu := \mathbb{E}[g(x)] \in \mathbb{R}^m$ における h の接平面を考えることで,

$$h(g(x)) \geq a^\top g(x) + b, \forall x \in \mathbb{R}^n, h(\mu) = a^\top \mu + b \quad (2.14)$$

を満たす $a \in \mathbb{R}^m$, $b \in \mathbb{R}$ がとれる。このとき,

$$\mathbb{E}[h(g(x))] \geq \mathbb{E}[a^\top g(x) + b] = a^\top \mu + b = h(\mu) \quad (2.15)$$

が成り立つ。 \square

定理 2.2.5 (ギブス (Gibbs) の不等式). $x \ll y$ を満たす任意の $x, y \in \text{rv}(\mathbb{R}^n)$ に対して $D_{\text{KL}}(x||y) \geq 0$ が成り立ち, 等号が成り立つのはこれらが同分布をもつ場合のみである。 \triangleleft

証明. 関数 $h(\cdot) := -\log(\cdot)$ の狭義凸性より, $g(x) = \varphi_y(x)/\varphi_x(x)$ として $D_{\text{KL}}(x||y) = \mathbb{E}[h(g(x))]$ にイエンセンの不等式を適用すると

^{†1} 確率密度関数 p, q に対しても $D_{\text{KL}}(p||q) := \int_{\mathbb{R}^n} p(x) \log \frac{p(x)}{q(x)} dx$ とする。

^{†2} i) $d(x, y) \geq 0$, ii) $d(x, y) = d(y, x)$, iii) $d(x, y) = 0 \Leftrightarrow x = y$, iv) $d(x, y) + d(y, z) \geq d(x, z)$

$$D_{\text{KL}}(x||y) \geq -\log \left(\int_{\mathbb{R}^n} \varphi_x(x) \frac{\varphi_y(x)}{\varphi_x(x)} dx \right) = -\log 1 = 0 \quad (2.16)$$

がしたがい、等号が成り立つののは $g(\cdot)$ が定数関数のときのみである。 \square

つぎに平均からの変位の類似度を量量化する。

定義 2.2.6 (分散共分散行列). $x \in \text{rv}(\mathbb{R}^n)$, $y \in \text{rv}(\mathbb{R}^m)$ に対して,

$$\text{Cov}[x, y] := \mathbb{E}[(x - \mathbb{E}[x])(y - \mathbb{E}[y])^\top] \in \mathbb{R}^{n \times m} \quad (2.17)$$

を x と y の**共分散行列** (covariance matrix),

$$\text{Var}[x] := \text{Cov}[x, x] \in \mathbb{R}^{n \times n} \quad (2.18)$$

を x の**分散行列** (variance matrix) と呼び、 $\text{Cov}[x, y] = 0$ のとき x と y は**無相関** (uncorrelated) であるという。 \triangleleft

定義と期待値演算の線形性から、

$$\text{Cov}[x, y] = \mathbb{E}[xy^\top] - \mathbb{E}[x]\mathbb{E}[y]^\top, \quad \text{Var}[x] = \mathbb{E}[xx^\top] - \mathbb{E}[x]\mathbb{E}[x]^\top \quad (2.19)$$

がしたがう。分散行列 $\text{Var}[x]$ は多次元確率変数 x の平均からのばらつきの度合いを量量化したものであり、その固有値・固有ベクトルにより、ばらつきの方向性を特徴づけることができる。すなわち、大きい固有値に属する固有ベクトルの方向ほど、ばらつきが大きい^{†1}。また、 x と y が独立ならば (2.19) より無相関であるが、その逆は成り立たない。つぎに、 $x - \mathbb{E}[x]$, $y - \mathbb{E}[y] \in \text{rv}(\mathbb{R})$ のなす角度 θ_{xy} に対応する量を導入して、類似度を量量化しよう。実は、 $\text{Cov}[x, y]$ は $x - \mathbb{E}[x]$ と $y - \mathbb{E}[y]$ の内積とみなせるため、共分散を規格化した**相関係数**

$$\rho(x, y) := \frac{\text{Cov}[x, y]}{\sqrt{\text{Var}[x]}\sqrt{\text{Var}[y]}} \in [-1, 1] \quad (2.20)$$

は $\cos \theta_{xy}$ と解釈できる（二つの多次元確率変数に対して、要素ごとの相関係数を並べた行列を**相関行列**と呼ぶ）。例えば、相関係数が 1 に近い場合、「 $x - \mathbb{E}[x]$ と $y - \mathbb{E}[y]$ は同じ方向を向いている」ため、 $(x - \mathbb{E}[x])/\sqrt{\text{Var}[x]}$ と $(y - \mathbb{E}[y])/\sqrt{\text{Var}[y]}$ の実現値は近い傾向がある^{†2}。

^{†1} 7.2.2 節の主成分分析を参照。

^{†2} 関連する話題として、8.2.1 節にて正準相関分析を紹介する。

2.2.2 多変量正規分布

本節では、代表的な確率分布モデルである正規分布に対して 2.2.1 節で導入したいくつかの統計量を具体的に計算することで、その特別な性質を整理する。

定理 2.2.7 (正規分布の統計量). \mathbb{R}^n -値確率変数 $x \sim \mathcal{N}(\mu, \Sigma)$ について、

$$\mathbb{E}[x] = \mu, \text{Var}[x] = \Sigma, \text{Mode}[x] = \mu \quad (2.21)$$

$$\Phi_x(\lambda) = \exp\left(j\lambda^\top \mu - \frac{1}{2}\lambda^\top \Sigma \lambda\right) \quad (2.22)$$

が成り立つ。 \triangleleft

注意 2.2.8 (退化正規分布). $x \sim \delta_\mu$ の特性関数は (2.1) より $\Phi_x(\lambda) = \exp(j\mu^\top \lambda)$ である。 $\Sigma \succ 0$ ではなく $\Sigma \succeq 0$ の場合、(2.22) を満たす x を退化した (degenerate) 正規分布と呼び、同様に $x \sim \mathcal{N}(\mu, \Sigma)$ と表記する。特に $\mathcal{N}(x|\mu, O) = \delta_\mu(x)$ は、正規分布の分散を小さくしていくとデルタ関数に収束することを示している。 \triangleleft

また、正規分布にしたがう確率変数の集合はアファイン変換に関して閉じている。

定理 2.2.9 (正規分布のアファイン変換). \mathbb{R}^n -値確率変数 $x \sim \mathcal{N}(\mu, \Sigma)$, $\bar{\mu} \in \mathbb{R}^q$, $A \in \mathbb{R}^{q \times n}$ に対して、

$$Ax + \bar{\mu} \sim \mathcal{N}(A\mu + \bar{\mu}, A\Sigma A^\top) \quad (2.23)$$

が成り立つ。 \triangleleft

さらに、和について閉じているという**再生性** (reproductive property) と呼ばれる性質は、線形システムの時間発展にともなう状態変数の分布の変化を見通しよく追跡することに役立つ。

定理 2.2.10 (正規分布の再生性). \mathbb{R}^n -値確率変数 $x \sim \mathcal{N}(\mu_x, \Sigma_x)$ と $y \sim \mathcal{N}(\mu_y, \Sigma_y)$ が独立ならば、 $x + y \sim \mathcal{N}(\mu_x + \mu_y, \Sigma_x + \Sigma_y)$ である。 \triangleleft

演習 2.1. (2.8), (2.9), (2.22) を用いて、定理 2.2.9, 2.2.10 を証明せよ。 \triangleleft

16 2. 確率と統計(19)

一般に, x, y が正規分布にしたがっても, $[x^\top y^\top]^\top$ が正規分布にしたがうとは限らない. ただし, x と y が独立であれば, つぎの結果がしたがう.

定理 2.2.11 (正規分布の独立性と無相関性). 正規分布にしたがう独立な確率変数 x と y に対して, $[x^\top y^\top]^\top$ は正規分布にしたがう. また, x と y が無相関かつ $[x^\top y^\top]^\top$ が正規分布にしたがうならば, x と y は独立な正規分布にしたがう. \triangleleft

証明. $z := [x^\top y^\top]^\top$, $z := [x^\top y^\top]^\top$ と表記する. $x \sim \mathcal{N}(\mu_x, \Sigma_x)$, $y \sim \mathcal{N}(\mu_y, \Sigma_y)$ が独立であるとすると,

$$\varphi_z(z) = \varphi_{(x,y)}(x, y) = \varphi_x(x)\varphi_y(y) = \mathcal{N}\left(\begin{bmatrix} x \\ y \end{bmatrix} \middle| \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}, \begin{bmatrix} \Sigma_x & O \\ O & \Sigma_y \end{bmatrix}\right)$$

であることから前半の主張がしたがう. つぎに

$$z \sim \mathcal{N}(\mu, \Sigma), \mu := \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}, \Sigma := \begin{bmatrix} \Sigma_x & \Sigma_{xy} \\ \Sigma_{xy}^\top & \Sigma_y \end{bmatrix} \quad (2.24)$$

と仮定する. このとき, 定理 2.2.7 および無相関性から $\Sigma_{xy} = \text{Cov}[x, y] = O$ である. ゆえに, $\varphi_z(z) = \mathcal{N}(z | \mu, \Sigma)$ が成り立つ. \square

2.3 統計的推論

再び x_A, x_B が「地点 A, B における風速」という物理量をモデル化した確率変数であるとし, さらに実際に地点 B における風速を観測したとする. このとき, この測定値を用いて地点 A における風速はどのように見積もることが妥当であろうか. 本節では, こうした観測による情報の取得を数学的に取り扱う統計的推論 (statistical inference) の枠組みを導入する.

2.3.1 ベイズ統計

(1) 情報取得の定式化 最初に, 本書におけるすべての議論の基盤となる条件付き確率を定義する.

定義 2.3.1 (条件付け). $x \in \text{rv}(X)$, $y \in \text{rv}(Y)$ および $y \in Y$ に対して,

$$\varphi_x(x|y=y) := \begin{cases} \frac{\varphi_{(x,y)}(x,y)}{\varphi_y(y)}, & \varphi_y(y) \neq 0 \\ 0, & \varphi_y(y) = 0 \end{cases} \quad (2.25)$$

を $y = y$ のもとでの x の条件付き確率密度関数 (conditional —) と呼ぶ. また, 条件付き確率, 条件付き期待値も同様に

$$\mathbb{P}_x(B|y=y) := \int_B \varphi_x(x|y=y) dx, \quad B \subset X \quad (2.26)$$

$$\mathbb{E}[x|y=y] := \int_X x \varphi_x(x|y=y) dx \quad (2.27)$$

と定義する. \triangleleft

記法 2.3.2 (条件付け). 本書では可読性に応じて, $\varphi_x(\cdot|y=y)$ を $\varphi_x|_{y=y}(\cdot)$, $\mathbb{P}_x(\cdot|y=y)$ を $\mathbb{P}_x|_{y=y}(\cdot)$ のように表記する. また, 条件付けの有無を強調する際は, φ_x を事前確率密度 (prior —), $\varphi_x|_{y=y}$ を事後確率密度 (posterior —) とも呼ぶ. \triangleleft

以下では, (2.25) からしたがう関係式

$$\varphi_{(x,y)}(x,y) = \varphi_y(y) \varphi_x|_{y=y}(x) \quad (2.28)$$

を踏まえて, $\varphi_x(x|y=y)$ をあたかも「 $y = y$ という情報 (観測結果) のもとでの $x = x$ の実現しやすさ」のように扱う. この解釈の妥当性はモデル化する現象と目的に即して判断されるべきであり本書の議論の対象とはしないが, 少なくともこれらの定義にもとづいて, 情報取得と呼ぶにふさわしい性質を有し, 様々な実用的な計算を可能にする, ベイズ統計学 (Bayesian statistics) と呼ばれる豊かな理論体系が構築できる. たとえば, 期待値の計算は各条件のもとでの個別の条件付き期待値の重み付き平均である.

定理 2.3.3 (塔特性). $x \in \text{rv}(X)$, $y \in \text{rv}(Y)$ に対して,

$$\mathbb{E}[x] = \int_Y \mathbb{E}[x|y=y] \varphi_y(y) dy \quad (2.29)$$

が成り立つ. \triangleleft

証明. (1.3) および (2.25) より

$$\varphi_x(x) = \int_Y \varphi_{(x,y)}(x,y) dy = \int_Y \varphi_x(x|y=y) \varphi_y(y) dy \quad (2.30)$$

である. これに x をかけて積分し, 積分順序を交換すると (2.29) を得る. \square

また, ベイズ統計は情報の逐次処理が正当化される理論体系となっている.

定理 2.3.4 (逐次推定の正当性). 確率変数 x, y, z および $x \in X, y \in Y, z \in Z$ が $\varphi_{(y,z)}(y,z) > 0$ を満たすとする. このとき, $z=z$ のもとでの事後分布 $\varphi_{(x,y)|z=z}$ を φ^z と表記すると,

$$\varphi_x(x|y=y, z=z) = \varphi_x^z(x|y=y) \quad (2.31)$$

が成り立つ. \triangleleft

証明. 定義 (2.25) を代入すれば確かめられる. \square

つまり, 無情報下で $(y,z) = (y,z)$ が測定された場合と, $z=z$ のもとで追加で $y=y$ が測定された場合では, 事後確率密度関数は一致する. これにより, 「 $z=z$ が得られた時点で, φ を φ^z に更新する」という手続きが正当化される. 再帰的に考えると, 時々刻々とデータが取得できる状況において, 蓄積したデータをまとめて利用するバッチ処理をおこなう必要はなく, 新たに取得したデータごとに逐次処理すればよいことを意味する.

こうした統計的推論においても正規分布は様々な有用な性質を有する. とりわけ, 事後分布が単純な代数計算により求まることは特に重要である.

定理 2.3.5 (結合正規分布の事後分布). 確率変数 x と y が

$$\begin{bmatrix} x \\ y \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}, \begin{bmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx}^\top & \Sigma_{yy} \end{bmatrix} \right) \quad (2.32)$$

を満たすとき,

$$\varphi_x|_{y=y} \sim \mathcal{N}(\hat{x}(y), \hat{\Sigma}), \quad (2.33)$$

$$\hat{x}(y) := \mu_x + \Sigma_{xy}\Sigma_{yy}^{-1}(y - \mu_y), \quad \hat{\Sigma} := \Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{xy}^\top \quad (2.34)$$

である。 \diamond

証明. 定義よりしたがう関係式 $\varphi_x|_{y=y}(x) \propto \varphi_{(x,y)}(x, y)$ および逆行列の公式

$$\begin{bmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{xy}^\top & \Sigma_{yy} \end{bmatrix}^{-1} = \begin{bmatrix} \hat{\Sigma}^{-1} & -\hat{\Sigma}^{-1}\Sigma_{xy}\Sigma_{yy}^{-1} \\ -\Sigma_{yy}^{-1}\Sigma_{xy}^\top\hat{\Sigma}^{-1} & * \end{bmatrix} \quad (2.35)$$

を用いて、直接計算によりしたがう。 \square

ここで、(x と y が無相関でなければ) $y = y$ という情報により x の不確実性が減少することに対応して、 $\hat{\Sigma} \preceq \Sigma_{xx}$ が成り立っている。また、 $\hat{\Sigma}$ が観測値 y に依存しないことに加え、推定誤差は観測値と独立となる。

定理 2.3.6 (正規分布における推定誤差の独立性). 定理 2.3.5 の記法のもと、

$$e := x - \hat{x}(y) \quad (2.36)$$

は y と $\hat{x}(y)$ のいずれとも独立である。 \diamond

演習 2.2. 後出の定理 2.3.7 を用いて、定理 2.3.6 を証明せよ。 \diamond

(2) 観測器としての確率変数 つぎに、観測器としての確率変数の価値の定量化方法にふれておく。そこで、確率変数 x の推定のために y を測定することの効果を定量化するとしよう。まず、 x と y が独立であれば y の観測値からは x に関する情報は得られないという事実はつぎのようにあらわされる。

定理 2.3.7 (独立性と事後分布). $x \perp\!\!\!\perp y$ であることと $\varphi_x|_{y=y}$ が $y \in Y$ に依存しないことは等価であり、このとき $\varphi_x|_{y=y} = \varphi_x$ である。 \diamond

証明. $x \perp\!\!\!\perp y$ と $\varphi_x|_{y=y} = \varphi_x, \forall y$ の等価性は (2.28) より自明。 $\varphi_x|_{y=y}$ が $y \in Y$ に依存しないとき、 $\varphi_x(x) = \int \varphi_x(x|y=y)\varphi_y(y)dy = \varphi_x(x|y=y)$ である。 \square

20 2. 確率と統計(19)

逆に独立でなければ何らかの情報が得られるが、その大きさはつぎのように定量化することができる。

定義 2.3.8 (相互情報量). 確率変数 x, y に対して,

$$\begin{aligned} I(x; y) &:= \int_Y D_{\text{KL}}(\varphi_x|_{y=y} \| \varphi_x) \varphi_y(y) dy \\ &= \int_Y \int_X \log \frac{\varphi_{(x,y)}(x, y)}{\varphi_x(x) \varphi_y(y)} \varphi_{(x,y)}(x, y) dx dy \end{aligned} \quad (2.37)$$

を相互情報量と呼ぶ。 \triangleleft

ここで、カルバック・ライブラ情報量は観測値 $y = y$ が得られた場合の x の密度関数の変化量をあらわしており、測定値 y に関する期待値をとることで確率変数 y に対する評価となっている。実際、 x と y が独立であるとき $I(x; y) = 0$ となり、これは情報が得られないことに対応する。

2.3.2 点推定

(1) 代表例 一般の事前分布に対して、事後確率分布を具体的に求めることは困難であることが多い、効率的かつ高精度に事後確率分布を求めることは、機械学習においても重要な課題の一つである。この事後確率分布をよく近似する確率分布（もしくは事後確率分布そのもの）を求めるのことを**分布推定**という^{†1}。たとえば、特定の確率分布モデルに限定し事後分布を最もよく近似するパラメータを求めるのことを全般的に**変分ベイズ法** (variational Bayesian methods) と呼ぶ。とりわけ、 $\hat{x}_{\text{PM}}(y) := \mathbb{E}[x|y=y]$ および $\hat{\Sigma} := \mathbb{E}[(x - \hat{x}_{\text{PM}}(y))(x - \hat{x}_{\text{PM}}(y))^T | y = y]$ を用いて正規分布 $\mathcal{N}(\hat{x}_{\text{PM}}(y), \hat{\Sigma})$ を事後分布 $x|_{y=y}$ の近似とする手法も、単純ではあるが広く用いられる^{†2}（図 2.3）。

一方で、確率変数 y の観測値 $y \in Y$ が与えられたとき、事後確率分布 $\varphi_x|_{y=y}$ を求める分布推定に対して、その代表値 $\hat{x} = \theta(y) \in X$ を与える（確率密度関数

^{†1} 事後分布の標本を生成する手法には、定理 10.2.8 のメトロポリス・ヘイスティングス法などがある。

^{†2} 5.2.3 節の非線形フィルタおよび 12 章のモーメントマッチングを参照。

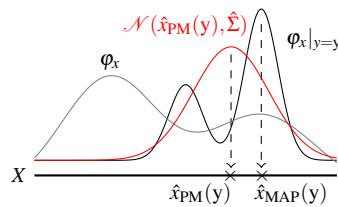


図 2.3: 分布推定と点推定

をデルタ関数に限定して分布推定する) ことを**点推定**という. 定義 2.3.8 では事後分布に大きな影響を与える (情報量の多い) 観測量の指標を与えたが, ここでは与えられた観測量を用いた点推定手法 θ を評価する. まずは, 点推定値 $\hat{x} := \theta(y) \in \sigma(y)$ も確率変数であることに注意し, いくつかの指標を導入する.

定義 2.3.9 (点推定値の評価). 確率変数 $x, \hat{x} \in \text{rv}(X)$ が,

$$\hat{x}_c(x) := \mathbb{E}[\hat{x}|x=x] = x, \forall x \in X \quad (2.38)$$

を満たすとき, \hat{x} は x の**不偏推定量** (unbiased estimator) であるという. また, 関数 $L : X \rightarrow \mathbb{R}_+$ に対して, $\mathbb{E}[L(x-\hat{x})]$ を損失関数 L に対する**ベイズリスク** (Bayes risk) と呼ぶ. \triangleleft

図 2.4 と図 2.5 に $\hat{x} \in \sigma(y)$ の場合における不偏性とベイズリスクの概念図を示す. 不偏性は, いかなる (真値) $x = x$ のもとでも期待値の意味では正しい推定値が得られることを意味し, 事前分布 φ_x には依存せず $\varphi_{y|x=x}$ のみで決まる性質である. 一方, L で定量化した誤差を事前分布も考慮して平均した値がベイズリスクであり, $\theta(y)$ が $\sigma(y)$ の中でベイズリスクを最小化するとき θ を**ベイズ推定関数** (Bayes estimator) と呼ぶ. 条件付き確率の解釈の妥当性を議論しないことは既に述べたが, 事後分布のいくつかの統計量を点推定値として用いることはベイズリスクの観点から正当化できる.

定理 2.3.10 (最小二乗誤差推定と事後平均). 事後平均 (posterior mean) $\hat{x}_{PM}(y) := \mathbb{E}[x|y=y]$ は二乗誤差 $L(e) := e^\top e$ に対するベイズ推定関数である. \triangleleft

証明. 定理 2.3.3 から, 任意の θ に対して,

$$\begin{aligned}\mathbb{E}[\|x - \theta(y)\|^2] &= \int_Y \mathbb{E}[\|x - \theta(y)\|^2 | y = y] \varphi_y(y) dy \\ &= \int_Y \left\{ \|\theta(y) - \hat{x}_{PM}(y)\|^2 + \underbrace{\mathbb{E}[x^\top x | y = y] - \|\hat{x}_{PM}(y)\|^2}_{\theta \text{に依存しない}} \right\} \varphi_y(y) dy\end{aligned}$$

が成り立つことから示される. \square

この定理は, 二乗誤差の意味で最適な点推定は事後分布の平均を用いることであると主張している. これに対して, 近似誤差が十分小さければ 0, それ以外は 1 と評価する場合, 事後分布の最頻値を用いることが最適となる.

定理 2.3.11 (MAP 推定). 最大事後確率 (maximum a posteriori; MAP) 推定

$$\hat{x}_{MAP}(y) := \arg \max_x \varphi_x|_{y=y}(x) \quad (2.39)$$

は十分小さい $\Delta > 0$ および一様損失関数

$$L(e) := 1 - \mathbb{1}_{(-\Delta, \Delta)}(e) \quad (2.40)$$

に対するベイズ推定関数である. \triangleleft

証明. 任意の θ に対して,

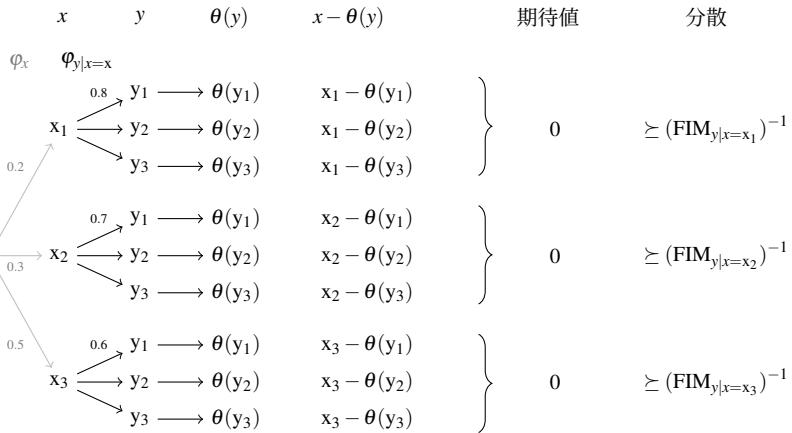
$$1 - \mathbb{E}[L(x - \theta(y)) | y = y] = \mathbb{P}(|x - \theta(y)| < \Delta | y = y) = \int_{\theta(y)-\Delta}^{\theta(y)+\Delta} \varphi_x|_{y=y}(x) dx$$

が成り立つことから示される. \square

ところで, 図 2.5において $y = y_1$ のもとで MAP 推定値は x_3 であるが, 素朴に $\varphi_{y|x=x}(y_1)$ を最大化する x である x_1 を採用することも, 観測結果を正当化する意味で妥当に思える. この考え方にもとづいた点推定値も広く用いられる.

定義 2.3.12 (尤度関数・最尤推定). $x \in \text{rv}(X), y \in \text{rv}(Y), y \in Y$ に対して,

$$l_x(x | y = y) := \varphi_y(y | x = x), x \in X \quad (2.41)$$

図 2.4: 観測値 y にもとづく x の推定値の不偏性 (事前分布 φ_x には依存しない)

を尤度関数 (likelihood function) と呼び、尤度関数もしくは等価的に対数尤度関数 $\log l_x(x|y=y)$ を最大化する

$$\hat{x}_{\text{ML}}(y) := \arg \max_{x \in X} \log l_x(x|y=y) \quad (2.42)$$

を最尤推定値 (maximum likelihood estimate) と呼ぶ。 \diamond

尤度関数 $l_x(\cdot|y=y)$ は非負値性は満たすが規格化はされていない。条件付き確率密度関数と尤度関数との関係、もしくは MAP 推定と最尤推定の違いは、事前情報 $\varphi_x(x)$ を掛け算の形で取り入れていることのみである。

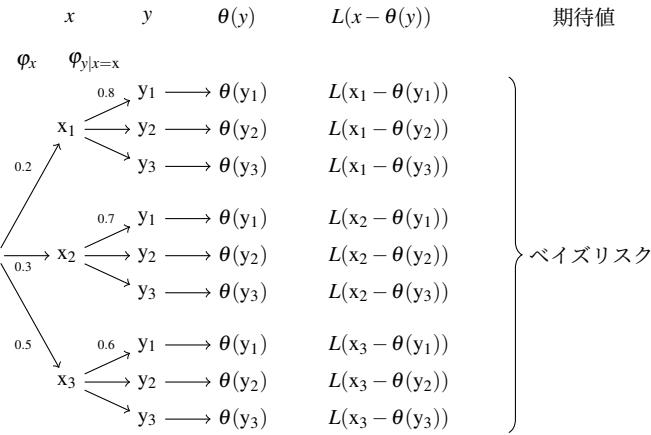
定理 2.3.13 (ベイズの定理)。定義 2.3.12 のもと、 $\varphi_y(y) \neq 0$ を満たす $y \in Y$ に対して、

$$\varphi_x(x|y=y) = \frac{l_x(x|y=y)\varphi_x(x)}{\varphi_y(y)} \quad (2.43)$$

が成り立つ。 \diamond

証明。 定義 2.3.1 よりただちにしたがう。 \square

(2) 精度の限界 三つの点推定方法を紹介したが、点推定性能の限界を

図 2.5: 観測値 y にもとづく x の推定値 $\theta(y) \in \sigma(y)$ のベイズリスク

特徴づけて、本章を締めくくる。準備として、最尤推定値は対数尤度関数の極値であることを踏まえて、対数尤度の勾配に関する二つの量を定義する。

定義 2.3.14. $x \in \text{rv}(\mathbb{R}^{n_x})$, $y \in \text{rv}(\mathbb{R}^{n_y})$ に対して,

$$\text{Sc}_x(x|y=y) := \frac{\partial}{\partial x} \log l_x(x|y=y) = \frac{1}{l_x(x|y=y)} \frac{\partial}{\partial x} l_x(x|y=y) \quad (2.44)$$

をスコア関数(score function)と呼び、

$$\text{FIM}_{y|x=x} := \int_Y \varphi_y(y|x=x) \text{Sc}_x(x|y=y) \text{Sc}_x(x|y=y)^\top dy \in \mathbb{R}^{n_x \times n_x} \quad (2.45)$$

をフィッシャー情報行列(Fisher information matrix)と呼ぶ。 \triangleleft

付録の補題 B.1 で示すように、フィッシャー情報行列は対数尤度関数のヘッセ行列を符号反転した行列に対応する。したがって、直感的には、この行列が大きいほど対数尤度関数は上にとがった形状をしており、その極大点である最尤推定値が見つけやすい。実際、このフィッシャー情報行列は、つきの定理の意味で「 x の真値が x の場合に、その推定のために y がもつ情報量」を特徴づける。

定理 2.3.15(クラメール・ラオ(Cramer-Rao)の不等式). $x \in \text{rv}(\mathbb{R}^{n_x})$, $y \in \text{rv}(\mathbb{R}^{n_y})$ および任意の $x \in \mathbb{R}^{n_x}$ に対して、 $\text{FIM}_{y|x=x}$ は正則であるとする。このとき、任意の x の不偏推定量 $\hat{x} \in \sigma(y)$ に対して

$$\text{Var}[\hat{x}|x=x] \succeq (\text{FIM}_{y|x=x})^{-1}, \forall x \in X \quad (2.46)$$

が成り立つ. \triangleleft

証明は付録 B.1 で与える. 真値が x のときの点推定誤差 ((2.19) を参照)

$$\mathbb{E}[(\hat{x} - x)(\hat{x} - x)^\top | x=x] = (\hat{x}_c(x) - x)(\hat{x}_c(x) - x)^\top + \text{Var}[\hat{x}|x=x] \quad (2.47)$$

において、「 $\text{FIM}_{y|x=x}$ が小さいとき, 第一項を 0 にするいかなる \hat{x} も第二項を小さくできない」ことを定理 2.3.15 は意味している. このように平均のずれと推定分散を同時に小さくすることは不可能であることを一般に, **バイアスとバリアンスのトレードオフ** と呼ぶ. また, (2.46) で等号が成り立つという意味で最適な不偏推定量を十分統計量 (sufficient statistics) と呼ぶ.

具体例として, 確率変数 x の要素を線形に重みづけた値 Cx を加法的雑音 w の存在下で観測する際のフィッシャー情報行列を考える.

定理 2.3.16 (正規性外乱下の線形観測におけるフィッシャー情報行列). \mathbb{R}^{n_x} -値確率変数 x と \mathbb{R}^{n_w} -値確率変数 $w \sim \mathcal{N}(\mu, \Sigma)$ が独立であるとする. ここで, $C^\top \Sigma^{-1} C$ が正則となる $C \in \mathbb{R}^{n_w \times n_x}$ に対して, $y := Cx + w$ を観測し x を推定する. このとき, $\text{FIM}_{y|x=x} = C^\top \Sigma^{-1} C$ であり,

$$\hat{x}(y) := (C^\top \Sigma^{-1} C)^{-1} C^\top \Sigma^{-1} (y - \mu) \quad (2.48)$$

は x の最尤推定値かつ十分統計量である. \triangleleft

証明. 定義より, 尤度関数は $l_x(x|y=y) = \varphi_{y|x=x}(y) = \varphi_w(y - Cx)$, 対数尤度は $\log l_x(x|y=y) = -(y - Cx - \mu)^\top \Sigma^{-1} (y - Cx - \mu)/2 + \text{定数}$, スコア関数は $\text{Sc}_x(x|y=y) = C^\top \Sigma^{-1} (y - Cx - \mu)$ である. したがって, スコア関数が 0 となる唯一の x である $\hat{x}(y)$ が最尤推定値である. さらに, $y|_{x=x} \sim \mathcal{N}(Cx + \mu, \Sigma)$ および (2.21) より $\text{FIM}_{y|x=x} = (C^\top \Sigma^{-1}) \text{Var}[y|x=x] (C^\top \Sigma^{-1})^\top = C^\top \Sigma^{-1} C$ がしたがう.

また, $A := (C^\top \Sigma^{-1} C)^{-1} C^\top \Sigma^{-1}$ とおくと $\hat{x}(y) = A(Cx + w - \mu) = A(w - \mu) + x$ である. 最後に, $x \perp\!\!\!\perp w$ に注意して, $x = x$ のもとで定理 2.2.9 を適用すると,

$$\hat{x}(y)|_{x=x} \sim \mathcal{N}(x, A\Sigma A^\top) = \mathcal{N}(x, (C^\top \Sigma^{-1} C)^{-1}) \quad (2.49)$$

が示され、求める結果が得られる。 \square

たとえば、別の推定値 $\theta'(y) := (C^\top C)^{-1} C^\top (y - \mu)$ は (2.38) の不偏性は満たすが、最小分散性 ((2.46) の等号) は成り立たない。

注意 2.3.17 (ガウス・マルコフ (Gauss-Markov) の定理)。定理 2.3.16において w の正規分布性を仮定しなくとも、 $\mathbb{E}[w] = \mu$, $\text{Var}[w] = \sigma^2 I$ (各要素が無相関かつ等分散) が成り立つならば、 $\theta'(y) := (C^\top C)^{-1} C^\top (y - \mu)$ はアファイン関数で与えられる不偏推定量の中では誤差分散を最小化示すことができ、線形最小分散不偏推定量 (best linear unbiased estimator; BLUE) と呼ばれる。

\diamond

記法 2.3.18. 記法の簡略化のため次章以降では、関係式において条件付けの実現値が任意であるとき、混乱を生じない限り省略する。たとえば、 $\mathbb{E}[x|y=y] = 0, \forall y$ は $\mathbb{E}[x|y] = 0$ と表記する。 \diamond

3 確率システム(17)

本章では、2章で紹介した確率を扱うための数学的な枠組みに時間の概念を導入し、システム制御理論における状態方程式のように因果性を有するシステムの解析をおこなう。また、線形制御理論の諸概念が、確率システムに対してどのように対応するかも整理する。

3.1 確率過程

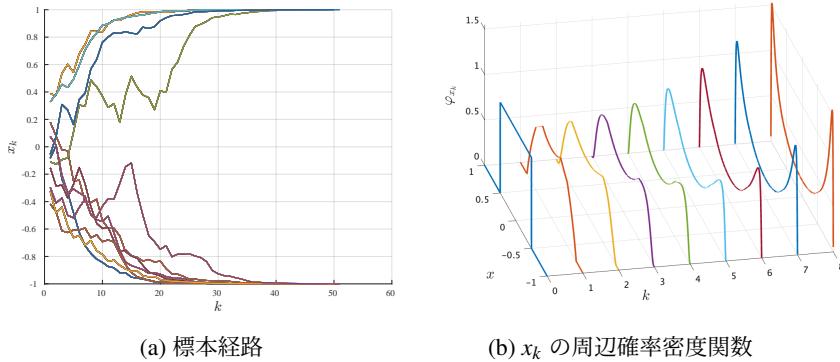
3.1.1 マルコフ過程

本書を通じて、時間変数 k が非負整数 $\mathbb{Z}_+ := \{0, 1, 2, \dots\}$ に値をとる離散時間の挙動を考察の対象とする。システム制御理論や信号処理においては、時々刻々と変化する量を $x(k)$ のような時間変数を引数にもつ関数[†]で表現し、 \mathbb{Z}_+ から X への関数を X -値（時系列）信号と呼んでいた。これに対して、確率的に変化する量の数学的な表現として、 \mathbb{Z}_+ から $\text{rv}(X)$ への関数を X -値の確率過程（stochastic process）と呼ぶ。慣例に従い、確率過程 x の時間変数は x_k のように下付きであらわす。また、特定の時間を抽出した組み合わせに対して

$$x_{k:l} = (x_k, x_{k+1}, \dots, x_l), \quad x_{k:} = (x_k, x_{k+1}, \dots) \quad (3.1)$$

などの記法を用い、 x が確率過程であることを明示する場合は $x_:$ と表記する。確率過程 $x_:$ は、確率変数を値にとる時系列信号として定義されているが、 $x_k \in \text{rv}(X)$

[†] 本書では、主に現代制御論の振り返りなどであらわれる、確率性をもたない信号の時間変数を括弧にてあらわす。

図 3.1: $[-1, 1]$ -値確率過程 x_k

に対して $(x_0, x_1, \dots) \in \text{rv}(X \times X \times \dots)$ (1.2 節) より時系列信号を値にとる確率変数とみなすこともでき、後者の視点での標本を**標本経路** (sample path) と呼ぶ。例として、区間 $[-1, 1]$ に値をとるある確率過程 x_k の 10 本の標本経路を図 3.1a に、 $k = 0, 1, \dots, 8$ に対する x_k の確率密度関数 φ_{x_k} を図 3.1b にそれぞれ示す。どちらにおいても、 k が大きくなるとともに、端点 (± 1) 付近に値が集中する様子が見てとれる。

現代制御論においては、状態変数と呼ばれる時系列信号を用いて、未来の状態が現在の状態にのみ依存し過去の状態に依存しない、あるいは現在の状態に過去の状態の情報が集約されている、という性質を表現していた[†]。物理量の時間変化を表現する際に自然にあらわれるこの因果的な性質を、ここでは確率過程の満たす条件として定式化する。「 x_k が状態変数である」ことの表現として、語感から $x_{k+1} \in \sigma(x_k)$ や $x_{k+1} \perp\!\!\!\perp x_{0:k-1}$ などを思い浮かべるかもしれない。しかし、どちらも適切ではなく、つぎの概念を用いて定義されるマルコフ性（定義 3.1.2）がこれに相当する。

定義 3.1.1 (条件付き独立)。 確率変数 x, y, z が

[†] 確率過程の統計的性質もしくはデータから状態変数を構成する方法は 8.2.1 節を参照。

$$\varphi_x|_{(y,z)=(y,z)} = \varphi_x|_{z=z}, \forall (y,z) \quad (3.2)$$

を満たすとき, x と y は z のもとで条件付き独立であるといい, $x \perp\!\!\!\perp_z y$ とあらわす. \triangleleft

ここで, x と y の独立性は, 任意の z に対して $x \perp\!\!\!\perp_z y$ であることを意味せず, その逆も成り立たない. 定理 2.3.4, 定理 2.3.7などを踏まえると, 条件付き独立 $x \perp\!\!\!\perp_z y$ は, 「 x に関して y からは z 以上の情報は得られない」, 「 y は z を通してのみ x に影響を与える」と解釈できる. これを用いて因果性を定義する.

定義 3.1.2 (マルコフ過程). 確率過程 x が

$$x_{k+1} \perp\!\!\!\perp_{x_k} x_{0:k-1}, \forall k \in \mathbb{Z}_+ \quad (3.3)$$

を満たすとき x をマルコフ過程 (Markov process) と呼び,

$$\Psi_k(x'|x) := \varphi_{x_{k+1}}(x'|x_k=x) \quad (3.4)$$

を x の遷移確率密度関数 (Transition —) と呼ぶ. \triangleleft

この定義に (3.2) の解釈を適用すると, マルコフ性は「状態の過去の実現値系列 $x_{0:k} = x_{0:k}$ のもとでの x_{k+1} の事後分布は $x_{0:k-1}$ に依存しない」ことを意味する. このマルコフ性を課すことで, 変数 k は単なる整数値パラメータではなくなる. なお, 条件付き確率密度関数の定義と (3.3) から, マルコフ過程 x の標本経路の確率密度関数は

$$\varphi_{x_{l:k}}(x_{l:k}) = \prod_{i=l}^{k-1} \Psi_i(x_{i+1}|x_i) \varphi_{x_l}(x_l) \quad (3.5)$$

で与えられる.

つぎに, 現代制御論で標準的な差分方程式による時間発展の表現として, 初期状態をあらわす確率変数 x_0 , 不確かな外生要因をあらわす確率過程 v , および関数 f を用いて,

$$x_{k+1} = f(x_k, v_k), k \in \mathbb{Z}_+ \quad (3.6)$$

30 3. 確率システム (17)

で定まる確率過程 x_i を考える。明らかに、 $x_{k+1} \in \sigma(x_k, v_k)$ であることから、 $x_k \in \sigma(x_0, v_{0:k-1})$, $k \geq 1$ である。このとき、 v_i に時間相関がなければ x_i はマルコフ過程となる。

仮定 3.1.3 (雑音の無記憶性). $v_{-1} := x_0$ の表記のもと、 $v_{-1:k-1} \perp\!\!\!\perp v_k$, $\forall k \in \mathbb{Z}_+$ である。 \triangleleft

とりわけ、 $\{v_k\}_{k \geq 0}$ が独立かつ同分布をもつとき、確率過程 v_i は**独立同分布** (i.i.d.: independently and identically distributed) にしたがうという。

定理 3.1.4 (確率システムのマルコフ性). 仮定 3.1.3 のもとで、(3.6) の x_i はマルコフ過程である。 \triangleleft

証明. (3.2) より、 $\varphi_{x_{k+1}}|_{x_{0:k}=x_{0:k}} = \varphi_{f(x_k, v_k)}|_{x_{0:k}=x_{0:k}}$ が $x_{0:k-1}$ に依存しないことを示せばよい。ここで、 $x_{0:k-1} \in \sigma(x_0, v_{0:k-2})$ に注意すると、仮定 3.1.3 と定理 1.2.5 より $x_{0:k-1} \perp\!\!\!\perp v_k$ であることから、求める結果がしたがう。 \square

ところで、確定システムの制御問題において代表的な制御目的は、状態変数を目標状態に収束させることであった。そこで、確率過程に対しても収束の概念を導入しておく。

定義 3.1.5 (確率過程の収束). \mathbb{R}^n -値確率過程 x_i および $\bar{x} \in \text{rv}(\mathbb{R}^n)$ に対して、

1. x_i が \bar{x} に**確率収束**する ($x_k \xrightarrow{p} \bar{x}$ と表記する) とは、任意の $\varepsilon > 0$ に対して、

$$\lim_{k \rightarrow \infty} \mathbb{P}(\|x_k - \bar{x}\| > \varepsilon) = 0$$

2. $p \geq 1$ に対して、 x_i が \bar{x} に p 次**平均収束**するとは、

$$\lim_{k \rightarrow \infty} \mathbb{E}[\|x_k - \bar{x}\|^p] = 0$$

3. x_i が \bar{x} に**概収束**するとは、

$$\mathbb{P}\left(\lim_{k \rightarrow \infty} x_k = \bar{x}\right) = 1$$

4. x_i が \bar{x} に法則収束するとは、任意の実数値有界連続関数 f に対して、

$$\lim_{k \rightarrow \infty} \mathbb{E}[f(x_k)] = \mathbb{E}[f(\bar{x})] \quad (3.7)$$

が成り立つことをいう。 \triangleleft

例えば、概収束は図 3.1a のような標本経路ごとの収束値、法則収束は図 3.1b のような確率密度関数の収束を定式化している。確率密度関数の収束である法則収束が他の三つの必要条件であることは、直感的に明らかであろう。また、証明は省略するが、概収束、 p 次平均収束はともに確率収束の十分条件である [7]。収束先が確率変数であるなど、時系列信号の収束よりも意味が捉えづらいため例をみておこう。

例 3.1.6 (確率過程の収束)。確率過程

$$x_{k+1} = 0.5x_k + v_k, x_0 \sim \mathcal{N}(0, 1) \quad (3.8)$$

を考える。まず、 v_i が $v_0 \sim \mathcal{N}(0, 1)$ のステップ信号 ($v_0 = v_1 = v_2 = \dots$) の場合の標本経路を図 3.2a に示す。このとき、すべての標本経路において $x_k - 2v_0 = 0.5^k \rightarrow 0$ が成り立ち、 x_k は $2v_0 \sim \mathcal{N}(0, 2^2)$ に概収束する。一方で、 $v_k \sim \mathcal{N}(0, 3)$, i.i.d. の場合の標本経路を図 3.2b に示す。後出の定理 3.2.4 で示すが、 $\varphi_{x_k}(\cdot)$ は $\mathcal{N}(\cdot | 0, 2^2)$ に法則収束する。

\triangleleft

演習 3.1. 例 3.1.6 の二つの場合に対する x_i がマルコフ過程か判定せよ。 \triangleleft

3.1.2 分布の時間発展

本節では、 X -値マルコフ過程 x_i に関して、その標本経路に注目するのではなく、図 3.1b のように周辺確率密度関数 $\{\varphi_{x_k}\}_{k \in \mathbb{Z}_+}$ がどのように時間とともに変化するか考察する。まず、確率密度関数の時間発展はつぎのように特徴づけられる。

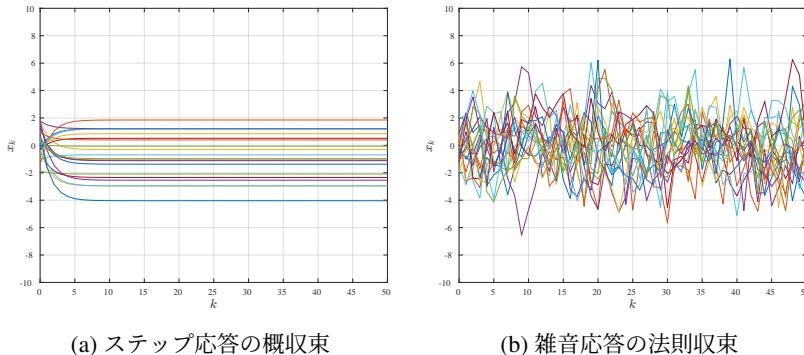


図 3.2: 確率過程の収束

定理 3.1.7 (チャップマン・コルモゴロフ (Chapman-Kolmogorov) の定理). (3.4) の遷移確率度関数をもつ \mathbb{R}^n -値マルコフ過程 x_i に対して,

$$\varphi_{x_{k+1}}(x') = \int_{\mathbb{R}^n} \Psi_k(x'|x) \varphi_{x_k}(x) dx \quad (3.9)$$

が成り立つ. \triangleleft

証明. 被積分関数は $\varphi_{(x_k, x_{k+1})}(x, x')$ であり, 積分により x_k が周辺化される. \square

例 3.1.8. 非線形確定システム

$$x_{k+1} = f_b(x_k) := x_k + 0.1(x_k - x_k^3) \quad (3.10)$$

を考える. 初期値 $x_0 \in [-1, 0]$ ならば $x_k \rightarrow -1$, $x_0 \in (0, 1]$ ならば $x_k \rightarrow 1$, という双安定性と呼ばれる性質を有する (図 3.3a). こうした性質は, 生体細胞のスイッチ機構や分化など多くの自然現象に現れる. 一方, 図 3.1 はこのシステムに雑音を加えた

$$x_{k+1} = f_b(x_k) + 0.5(1 - |x_k|)v_k, \quad v_k \sim \text{Uni}([-1, 1]) \text{ i.i.d.} \quad (3.11)$$

で与えられるマルコフ過程 x_i である. 雜音の大きさは, (3.10) の安定平衡点 $x = \pm 1$ では 0 で, 二点の中点近くでは大きい. その結果, 図 3.3b に示す $x_0 = -0.5$

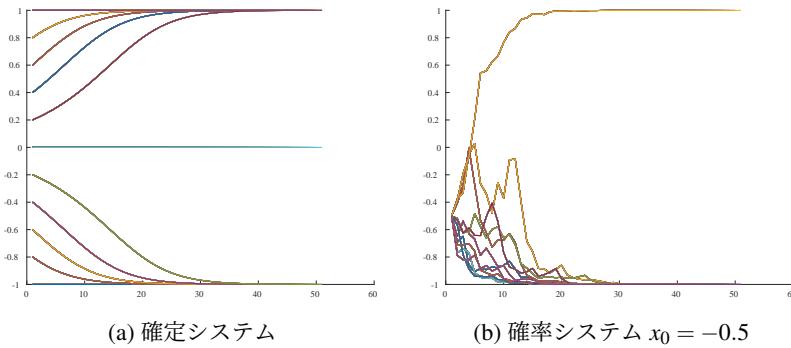


図 3.3: 双安定性システム

に対する標本経路においても、収束先が初期値だけでは決まらない。また、図 3.1a は $x_0 \sim \text{Uni}([-0.5, 0.5])$ に対する標本経路、図 3.1b は $\varphi_{x_0} = \mathbb{1}_{[-0.5, 0.5]}$ に対して (3.9) を再帰的に計算した結果であり、 $\varphi_{x_\infty} = (\delta_1 + \delta_{-1})/2$ に近づく。△

演習 3.2. 例 3.1.8 の確率システムの遷移確率密度関数を求めよ。△

注意 3.1.9 (分布の時間発展と経路の分布)。分布の時間発展の情報だけからは、異なる時刻間の結合確率密度関数はわからないため、標本経路も構成できない。たとえば、図 3.1 の例において、実際には図 3.1a のように各標本経路は収束している。しかしながら、図 3.1b の $\{\varphi_{x_k}\}_{k \in \mathbb{Z}_+}$ の情報だけからは、経路の確率密度関数 $\varphi_{x_{0:10}}$ はわからないため、 $x_k = (-1)^k$ や $x_k = (-1)^{k+1}$ のような振動的な標本経路が多数含まれる可能性も否定できない。この情報の欠落を受け入れるかわりに確率を含まない線形の時間発展表現を得たのが定理 3.1.7 といえる。△

3.1.3 リアブノフ安定論

一般的に \mathbb{R}^n -値時系列信号 $x(k)$ の振る舞いを、視覚的に理解することや直接的に解析することは容易ではない。これに対して、リアブノフ関数と呼ばれる実数値関数 $V: \mathbb{R}^n \rightarrow \mathbb{R}$ を導入し、 $V(x(k))$, $k \in \mathbb{Z}_+$ の挙動を調べることで、間接的に $x(k)$ の挙動を特徴づけることができる。例えば、常に $V(x(k))$ が k に関

して非増加であるならば、レベルセット

$$\text{Lv}(V; r) := \{x \in \mathbb{R}^n : V(x) \leq r\} \quad (3.12)$$

の記法のもと、 $x(k)$ は $\text{Lv}(V; V(x(0)))$ にとどまり続ける。このような方法論は **リアブノフ安定論** (Lyapunov theory) と呼ばれ、確定システムに対するシステム制御理論や力学系理論において様々な証明手法・実用的計算手法を提供してきた [8]。ここでは、確率システムに対して同様の考察をおこなう。

そこで、 X -値マルコフ過程 x と実数値関数 $V : \mathbb{Z}_+ \times X \rightarrow \mathbb{R}$ に対して、実数値確率過程 $V(k, x_k)$, $k \in \mathbb{Z}_+$ を考える。ここで、期待値の意味での関数 V の x に沿った時間変化量を

$$(\mathcal{L}_x[V])(k, x) := \mathbb{E}[V(k+1, x_{k+1})|x_k = x] - V(k, x) \quad (3.13)$$

と定義すると、 $V(k, x_k)$ の期待値の時間変化はつきの公式で求まる。

定理 3.1.10 (ディンキン (Dynkin) の公式). X -値マルコフ過程 x および $V : \mathbb{Z}_+ \times X \rightarrow \mathbb{R}$ に対して、 $V(k, x_k)$ の期待値は

$$\mathbb{E}[V(k, x_k)] = \mathbb{E}[V(0, x_0)] + \sum_{i=0}^{k-1} \mathbb{E}[\mathcal{L}_x[V](i, x_i)] \quad (3.14)$$

を満たす。 \triangleleft

証明. (3.13) の両辺に $\varphi_{x_k}(x)$ を掛けて x に関して積分すると、

$$\mathbb{E}[\mathcal{L}_x[V](k, x_k)] = \mathbb{E}[V(k+1, x_{k+1})] - \mathbb{E}[V(k, x_k)], \forall k \in \mathbb{Z}_+$$

であることからしたがう。ただし、右辺第一項に塔特性 (2.29) を用いた。 \square

したがって、 $\mathcal{L}_x[V]$ が 0 以下の値のみを取る関数ならば、 $V(k, x_k)$ の期待値は非増加である。

注意 3.1.11. 定理 3.1.10 は各時刻における $V(k, x_k)$ の分布に対する結果でしかなく、個別の軌道の性質を直接特徴づけるものではない。例えば、 $\mathcal{L}_x[V](k, x)$

が任意の k, x に対して非正であっても, $V(k, x_k)$ の軌道は一般に単調非増加となる. 実際, 例 3.1.8において, -1 からの距離 $V_-(x) := 1 + x$ を考えると,

$$\mathcal{L}[V_-](k, x) = 0.1(x - x^3) \leq 0, \forall x \in [-1, 0] \quad (3.15)$$

であるが, レベルセット $L_v(V_-, 1)$ から脱出する軌道が存在する (図 3.3b). この脱出確率なども, マルチングール (martingale) 不等式 (付録 A.3)などを用いてある程度は評価できる. \triangleleft

この定理の利用法の一つとして, $V(x_k)$ の期待値が指数的に 0 に収束するための十分条件を与えておこう.

系 3.1.12. X -値マルコフ過程 $x, V : X \rightarrow \mathbb{R}, \beta > 0$ に対して,

$$\mathbb{E}[V(x_{k+1})|x_k = x] \leq \beta V(x), \forall k, x \quad (3.16)$$

ならば,

$$\mathbb{E}[V(x_k)] \leq \beta^k \mathbb{E}[V(x_0)] \quad (3.17)$$

が成り立つ. \triangleleft

証明. $V'(k, x) := \beta^{-k}V(x)$ とおくと, $\mathcal{L}_x[V'](k, \cdot) \leq 0$ である. これに定理 3.1.10 を適用すると, $\mathbb{E}[V'(k, x_k)] \leq \mathbb{E}[V'(0, x_0)]$ が成り立つ. \square

たとえば $\beta \in (0, 1), V(x) = \|x - \bar{x}\|^p$ に対して (3.16) が成り立つならば, x_k は \bar{x} に p 次平均収束することなどが証明できる. 他にも, $\mathbb{E}[|x_k|^p]$ の上限が求まれば, つぎの定理を用いて分布の裾の重さ $\mathbb{P}(|x_k| \geq a)$ を評価することもできる[†].

定理 3.1.13. 1. (マルコフ (Markov) の不等式) 非負値確率変数 x が有限の期待値 $\mathbb{E}[x] < \infty$ をもつとき, 任意の $a > 0$ に対して,

$$\mathbb{P}(x \geq a) \leq \frac{1}{a} \mathbb{E}[x] \quad (3.18)$$

が成り立つ.

[†] 凸最適化を用いて (3.16) などを判定する手法や一般的な多項式の期待値を見積もる手法を 11.1 節で紹介する.

36 3. 確率システム (17)

2. (チェビシェフ (Chebyshev) の不等式) 実数値確率変数 x と任意の $a > 0, p > 0$ に対して,

$$\mathbb{P}(|x| \geq a) \leq \frac{1}{a^p} \mathbb{E}[|x|^p] \quad (3.19)$$

が成り立つ.

△

証明. 非負値性に注意して,

$$\int_0^\infty x \varphi_x(x) dx \geq \int_a^\infty x \varphi_x(x) dx \geq a \int_a^\infty \varphi_x(x) dx = a \mathbb{P}(x \geq a) \quad (3.20)$$

より (3.18) がしたがう. また, (3.18) を用いて

$$\mathbb{P}(|x| \geq a) = \mathbb{P}(|x|^p \geq a^p) \leq \frac{\mathbb{E}[|x|^p]}{a^p} \quad (3.21)$$

が成り立つ. □

3.2 線形確率システム解析

本節では, $A \in \mathbb{R}^{n_x \times n_x}$, $B \in \mathbb{R}^{n_x \times n_v}$, $C \in \mathbb{R}^{n_y \times n_x}$ を用いた線形の差分方程式

$$x_{k+1} = Ax_k + Bv_k, \quad (3.22)$$

$$y_k = Cx_k \quad (3.23)$$

により与えられる確率過程に対する基礎的な結果を紹介する.

3.2.1 時間領域

まず, (3.22) で定まる x_k は畳み込みの形で

$$x_k = A^k x_0 + \sum_{i=0}^{k-1} A^{k-1-i} B v_i, \quad k \geq 1 \quad (3.24)$$

と陽に書き下せることから, 期待値と分散は以下の通り時間発展する.

定理 3.2.1 (線形確率システムの期待値・分散の時間発展). (3.22)において,

$$\mathbb{E}[x_{k+1}] = A\mathbb{E}[x_k] + B\mathbb{E}[v_k] \quad (3.25)$$

が成り立つ. また, $x_0, \{v_k\}_{k \geq 0}$ が無相関であるとき,

$$\text{Var}[x_{k+1}] = A\text{Var}[x_k]A^\top + B\text{Var}[v_k]B^\top \quad (3.26)$$

である. \diamond

この定理もしくは (3.24) から

$$\mathbb{E}[x_k] = A^k \mathbb{E}[x_0] + \sum_{i=0}^{k-1} A^{k-1-i} B \mathbb{E}[v_i], \quad k \geq 1 \quad (3.27)$$

$$\text{Var}[x_k] = A^k \text{Var}[x_0] A^{k^\top} + \sum_{i=0}^{k-1} A^i B \text{Var}[v_{k-1-i}] (A^i B)^\top \quad (3.28)$$

が成り立つ. こうした平均と分散に関して閉じた漸化式が得られるのは, システムの線形性の恩恵である. さらに, 仮定 3.1.3 のもと $x_0, v_k, k \in \mathbb{Z}_+$ がすべて正規分布にしたがうとき, (3.22) に定理 2.2.9 と定理 2.2.10 を再帰的に適用すれば, $x_k \in \sigma(x_0, v_{0:k-1})$ は v_k と独立な正規分布となる. この性質は, 5 章のフィルタの導出において中心的な役割を果たす.

つぎに, 現代制御論における可制御性の観点から, これらの結果を眺めてみよう. 確定システム

$$x(k+1) = Ax(k) + Bu(k), \quad x(0) = x_0 \quad (3.29)$$

の一般解は可制御性行列 C_k を用いて

$$x(k) = A^k x_0 + C_k \begin{bmatrix} u(k-1)^\top & \cdots & u(0)^\top \end{bmatrix}^\top \quad (3.30)$$

$$C_k := \begin{bmatrix} B & BA & \cdots & BA^{k-1} \end{bmatrix}, \quad (3.31)$$

と書ける. ここで, (3.30) から, 「任意の $x \in \mathbb{R}^{n_x}$ に対して $x(k) = x$ を実現する入力が存在する」(時刻 k においてシステムは可到達という) ことと C_k の行フルランク性, もしくは可到達性グラミアン

$$G_k^r := C_k C_k^\top = \sum_{i=0}^{k-1} A^i B B^\top (A^\top)^i \quad (3.32)$$

の正則性は等価である^{†1}. この行列は, $\text{Var}[v_k] = I$, $k \in \mathbb{Z}_+$ のときの (3.28) の右辺第二項に一致する. つまり, $x_0 = x_0$, $v_k \sim \mathcal{N}(0, I)$, $\forall k \in \mathbb{Z}_+$ のとき $x_k - A^k x_0$ は共分散行列 G_k^r をもつ (非退化な) 正規分布であるため, $\phi_{x_k}(x) > 0$, $\forall x$ である. したがって, 確定システムの場合の可到達性は, 確率システムにおいて「各入力チャネルに独立な正規分布にしたがう雑音が印加されると, 任意の状態 x 近傍への遷移が正の確率で実現する」ことに対応している. こうした性質は強化学習における探索雑音の性質を調べる際にも重要である^{†2}.

最後に, 安定性の違いを簡単に確認する. 入力を $u(k) = 0$ とした (3.29) の確定システムの状態 $x(k)$ が初期状態に依存せず原点に収束するための必要十分条件は, A のスペクトル半径 (固有値の絶対値の最大値) $\rho(A)$ が $\rho(A) < 1$ (A がシュア安定 (Schur stable) であるという) を満たすことである. 同様に, (3.25) からつぎの結果が成り立つ.

定理 3.2.2 (線形システムにおける状態の期待値の漸近安定性). (3.22)において $\mathbb{E}[v_k] = 0$ とする. このとき, $\mathbb{E}[x_0]$ が有界となる任意の x_0 に対して $\lim_{k \rightarrow \infty} \mathbb{E}[x_k] = 0$ となるための必要十分条件は $\rho(A) < 1$ である. \triangleleft

証明. $\mathbb{E}[x_k] = A^k \mathbb{E}[x_0]$ より自明. \square

また, 現代制御論においてリアプノフ方程式の正定解の存在性とシステムの安定性の等価性 [1, 定理 5.6] は, 重要な役割を果たしていた. これに対応する結果はつぎの定理で与えられる.

定理 3.2.3 (安定性とリアプノフ方程式). $\rho(A) < 1$ ならば, 任意の $Q \succeq O$ に対して,

^{†1} ケーリー・ハミルトンの定理より $k \geq n_x$ に対して $\text{rank } G_k^r = \text{rank } G_{n_x}^r$ である.

^{†2} さらに??節において非線形システムに対する最小エネルギー制御則の観点から興味深い結果を導出する.

$$AGA^\top + Q = G \quad (3.33)$$

$$A^\top \Pi A + Q = \Pi \quad (3.34)$$

はともに唯一解 \bar{G} , $\bar{\Pi}$ をもち, $\bar{G} \succeq Q$, $\bar{\Pi} \succeq Q$ を満たす. 逆に, ある $Q \succ O$ に対して (3.33) もしくは (3.34) が正定解をもてば, $\rho(A) < 1$ である. \triangleleft

証明. $\rho(A) = \rho(A^\top)$ より, (3.34) の場合のみを示せば十分. 安定性より

$$\bar{\Pi} := \sum_{k=0}^{\infty} (A^\top)^k Q A^k \succeq Q \quad (3.35)$$

は有限確定し, (3.34) を満たすことは明らか. また, 任意の解 Π に対して,

$$\bar{\Pi} = \sum_{k=0}^{\infty} (A^\top)^k (\underbrace{(\Pi - A^\top \Pi A)}_{=Q}) A^k = \Pi - \lim_{k \rightarrow \infty} \{(A^\top)^k \Pi A^k\} = \Pi \quad (3.36)$$

より一意性がしたがう. 後半に関しては, $V(k) := x_i^\top (A^\top)^k \Pi A^k x_i \geq 0$ とおくと,

$$V(k+1) - V(k) = -x_i^\top (A^\top)^k Q A^k x_i \leq -\sigma_{\min}(Q) \|A^k x_i\|^2 \leq 0 \quad (3.37)$$

である ($\sigma_{\min}(Q) > 0$ は Q の最小特異値). したがって, $V(k)$ は下に有界な単調非増加列であるため収束値をもつ. よって $V(k+1) - V(k) \rightarrow 0$ より, 任意の x_i に対して $\|A^k x_i\| \rightarrow 0$ が成り立つ. これは $\rho(A) < 1$ を意味する. \square

ここで, (3.34) の一意解 $\bar{\Pi}$ は $v_i = 0$ のもとで初期状態 x_i が状態過程 $x_k = A^k x_i$ の大きさに与える影響を

$$\sum_k \|x_k\|_Q^2 = x_i^\top \bar{\Pi} x_i \quad (3.38)$$

により特徴づける. これに対して, (3.33) の一意解 $\bar{G} := \sum_k A^k Q (A^\top)^k$ は定常外乱 v_i が状態 x_i の定常分散に与える影響を特徴づける. 具体的には, $Q = BB^\top$ とおいて (3.32) と比較すると, シュア安定な A に対して

$$G_\infty^r := \lim_{k \rightarrow \infty} G_k^r \quad (3.39)$$

は,

$$AGA^\top + BB^\top = G \quad (3.40)$$

の唯一解である。この特性は (3.26) の不動点とみることもでき、確率過程の収束として記述するところの結果となる（例 3.1.6）。

定理 3.2.4 (線形システムの定常分散). 線形システム (3.22)において, A はシーア安定, 独立同分布の $v_k \sim \mathcal{N}(0, I)$ は x_0 と独立とする。このとき, x_k は $\mathcal{N}(0, G_\infty^r)$ に法則収束する。 \triangleleft

証明. $x_k - A^k x_0 \sim \mathcal{N}(0, G_k^r)$ よりしたがう。厳密な証明および正規分布を仮定しない場合の結果に関しては [9, Lemma 3.2]などを参照。 \square

安定な確率システムが定常的にどのように揺らぐか調べることは工学的にも重要であり、線形システムの場合は (3.40) を解くことで $\mathcal{N}(0, G_\infty^r)$ と求まる。

3.2.2 周波数領域

ここでは、確率過程の特性を周波数領域で解析する結果をいくつか紹介しておこう。まず、連続時間信号のラプラス変換と同様に、ベクトルもしくは行列の列 $\{x(k)\}_{k \in \mathbb{Z}_+}$ に対して（ベクトル値もしくは行列値）複素関数

$$Z_x(z) := \sum_{k \geq 0} z^{-k} x(k), z \in \mathbb{C} \quad (3.41)$$

を x の z 変換と定義する[†]。また、絶対値の和が有界な実数値信号 $x(k)$, $k \in \mathbb{Z}_+$ および $\omega \in [0, 2\pi]$ に対して、 $Z_x(e^{j\omega})$ を離散フーリエ変換と呼ぶ。これは、 $x(k) = (1/2\pi) \int_0^{2\pi} Z_x(e^{j\omega}) e^{j\omega k} d\omega$ を満たすため信号 x に含まれる周波数 ω 成分と解釈できる。つぎに確率過程 x に対しても周波数の概念を

$$S_x(\omega) := \lim_{\bar{k} \rightarrow \infty} \frac{1}{\bar{k}} \mathbb{E} \left[\left| \sum_{k=0}^{\bar{k}} x_k e^{-j\omega k} \right|^2 \right] \quad (3.42)$$

[†] 正確には、収束半径内においてはこの級数により定義し、その外へ解析接続する。

により導入し、**パワースペクトル密度** (Power Spectral Density) と呼ぶ。標本経路ごとの離散フーリエ変換の絶対値の二乗のみに注目し、その期待値が(3.42)であるとみなせるが、長時間平均を考えることで図3.2bのような時間的に減衰しない確率過程にも適用できる点に注意する。

つぎに、線形システム $x(k+1) = Ax(k) + Bv(k)$, $y(k) = Cx(k)$ において

$$g(0) = O, g(k) := CA^{k-1}B, k \geq 1 \quad (3.43)$$

をインパルス応答、その z 変換

$$G(z) := C(zI - A)^{-1}B \quad (3.44)$$

を**伝達関数行列** (transfer matrix) と呼ぶ。この伝達関数行列の分母多項式の根は極と呼ばれ、分子との相殺がなければ A の固有値に一致する。また、 $x(0) = 0$ のとき、

$$Z_y(z) = G(z)Z_v(z) \quad (3.45)$$

が成り立つ ((3.24) の表現から確認できる)。

最後に、安定なシステムに正弦波を印加した場合の出力 (周波数応答と呼ばれる) を考えよう。 (3.45) に $z = e^{j\omega}$ を代入すると、入力の各周波数成分が $G(e^{j\omega})$ 倍されて出力されることがわかる。伝達関数によって周波数ごとのゲインが規定されるという古典制御理論 [10] の核となるこの性質は、確率システムに対してはつぎの定理で特徴づけられる。

定理 3.2.5 (雑音の白色性)。 実数値確率過程 v_k が無相関かつ $\mathbb{E}[v_k] = 0$, $\text{Var}[v_k] = 1$, $\forall k \in \mathbb{Z}_+$ を満たすとき, $S_v(\omega) \equiv 1$ である。また、線形システム $\bar{x}_{k+1} = A_w \bar{x}_k + B_w v_k$, $\bar{v}_k = C_w \bar{x}_k$ が一入力一出力で安定であるとき、出力過程 \bar{v} は

$$S_{\bar{v}}(\omega) = |G_w(e^{j\omega})|^2 \quad (3.46)$$

を満たす。ただし、 $G_w(z) := C_w(zI - A_w)^{-1}B_w$ である。 \triangleleft

証明. 古典制御理論における、時間領域の畳み込み積が周波数領域の乗算となることの証明と同様に示される。具体的には、

$$\begin{aligned} \sum_{k=0}^{\bar{k}} \bar{v}_k e^{-j\varpi k} &= \sum_{k=0}^{\bar{k}} \sum_{i=0}^k g(k-i) v_i e^{-j\varpi k} \quad (g \text{ は (3.43) と同様に定義}) \\ &= \sum_{i=0}^{K-1} \sum_{k'=0}^{\bar{k}-i} g(k') v_i e^{-j\varpi i} e^{-j\varpi k'} \quad (k' := k - i) \\ &= \sum_{i=0}^{\bar{k}-1} v_i e^{-j\varpi i} \sum_{k'=0}^{K-i} g(k') e^{-j\varpi k'} \end{aligned}$$

を $S_{\bar{v}}$ の定義に代入すればよい。 \square

この v のように全周波数帯域に同一の強度をもつという意味で、各時刻で平均 0 の独立同分布をもつ確率過程は**白色雑音** (white noise) と呼び、それ以外を**有色雑音** (colored noise) と呼ぶ。白色雑音を「周波数重み」の役割を果たす安定なシステムに通すことで有色雑音を生成することは、確率的な時系列現象のモデリングで基礎的な考え方である。

例 3.2.6 (有色雑音). 天候の不確実性により変動する風力発電量の時間変化をモデリングする際、物理法則（タービンの慣性が大きく、発電量は高周波振動しないなど）や過去の実績から図 3.4a の破線のような周波数特性をもつことがわかっているとする。これに対して、図 3.4a の実線の周波数ゲインをもつ A_w, B_w, C_w を用意することは難しくない ($\varpi = 0.5$ をカットオフ周波数とする 3 次のチェビシェフフィルタを用いた)。このとき、 v を白色雑音として、 $\bar{x}_{k+1} = A_w \bar{x}_k + B_w v_k$, $\bar{v}_k = C_w \bar{x}_k$ の出力 \bar{v}_k を時刻 k における風力発電量モデルに用いることは理にかなっている。実際、図 3.4b に示す \bar{v} の標本経路は白色雑音よりも滑らかであり、高周波成分は限定的であるという事前情報が反映できている。 \diamond

なお、例 3.2.6 の \bar{v} はマルコフ性や無記憶性を満たさない。次章以降は主に外乱の白色性を想定して理論を構築するが、有色雑音 \bar{v} により駆動されるシス

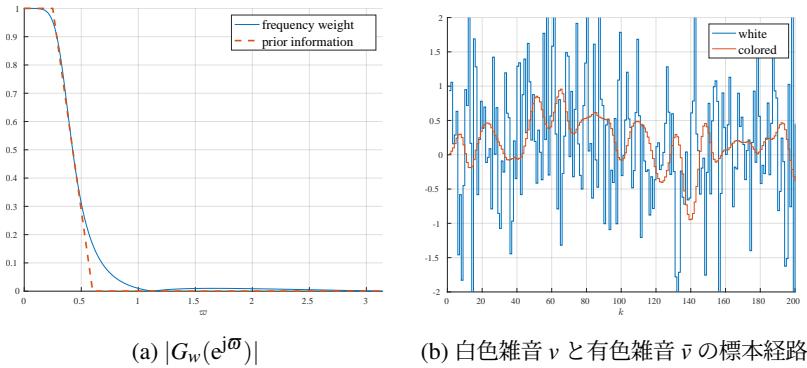
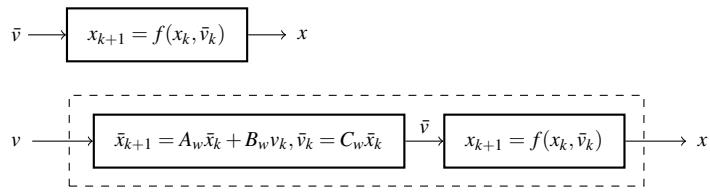


図 3.4: 有色雑音

図 3.5: 有色雑音 \bar{v} に駆動されるシステム (v は白色雑音)

システム $x_{k+1} = f(x_k, \bar{v}_k)$ を考える場合も、周波数重みを入力端に付加することで白色雑音に駆動されるシステムとみなせる（図 3.5）。

4 確率制御 1 (18)

本章では、入力を有する確率システムを導入し、最適制御問題の定式化および解法を概説する。確定システムの場合と比較して、制御則の因果性には注意する必要があるが、導出の手順はほぼ同様である。一方で、現代制御論における主結果である二次形式のコスト関数のもとでの線形システムの最適制御則は、確率的な雑音の影響下においても最適であるなど、自明ではない結果も成り立つ。

4.1 動的計画法

4.1.1 因果性とフィードバック制御

(離散時間の) システム制御理論では、与えられた関数 f および初期条件 $x(0)$ に対して、差分方程式

$$x(k+1) = f(x(k), u(k)), k \in \mathbb{Z}_+ \quad (4.1)$$

により定まる入力信号 $u(k)$ から状態変数信号 $x(k)$ への対応を (ダイナミカル) システムと呼んでいた。同様に、関数 $f_k : \mathbb{X} \times \mathbb{U} \times \mathbb{V} \rightarrow \mathbb{X}$, $k \in \mathbb{Z}_+$, および仮定 3.1.3 を満たす初期状態をあらわす $x_0 \in \text{rv}(\mathbb{X})$, 雜音 (不確かな外生信号) をあらわす \mathbb{V} -値確率過程 v に対して

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

により定まる、 \mathbb{U} -値の入力確率過程 u から \mathbb{X} -値の状態確率過程 x への対応を 確率システム (stochastic system) と呼ぶ。また、特定の入力 u に対応する状態過程であることを明記する場合は x^u と表記する。さらに、入力が

$$u_{0:k} \perp\!\!\!\perp v_{k+1}, \forall k \in \mathbb{Z}_+ \quad (4.3)$$

という意味で未来の雑音に影響を与えない（因果的という）ならば

$$(x_{0:k}^u, u_{0:k}) \perp\!\!\!\perp v_{k+1} \quad (4.4)$$

である。したがって、 f_k, v_k のみで定まる遷移確率密度関数と呼ばれる関数 Ψ_k が存在し、任意の因果的な u に対して

$$\varphi_{x_{k+1}^u}(x' | x_{0:k}^u = x_{0:k}, u_{0:k} = u_{0:k}) = \Psi_k(x' | x_k, u_k), \forall k, x_{0:k}, u_{0:k} \quad (4.5)$$

を満たす。たとえば、 $v_k \sim \mathcal{N}(0, \sigma_k^2 I)$, i.i.d. および,

$$x_{k+1} = f'(x_k, u_k) + g(x_k, u_k)v_k, k \in \mathbb{Z}_+ \quad (4.6)$$

により与えられる確率システムの遷移確率密度関数は、

$$\Psi_k(x' | x, u) = \mathcal{N}(x' | f'(x, u), \sigma_k^2 g(x, u)g(x, u)^\top) \quad (4.7)$$

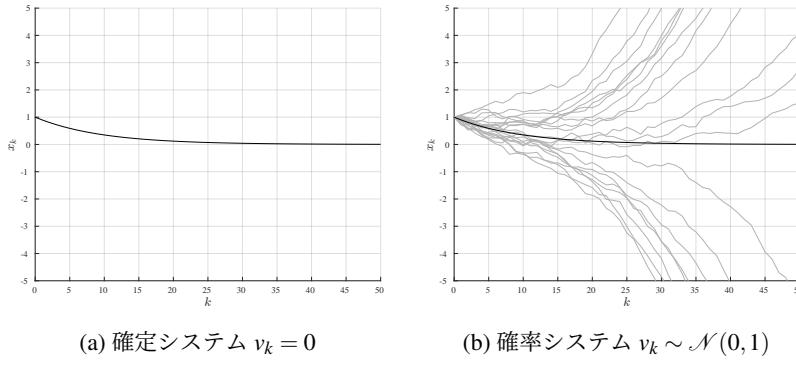
で与えられる。ここからの議論では、(4.2) だけでなく Ψ_k を直接的に用いることも多い。

一般に、状態過程 x^u が望ましい性質を有するように u を定める問題を制御問題と呼ぶ。まず、確定システムと確率システムの制御問題を比較するために、つぎの例を考える。

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

$$x(k+1) = ax(k) + u(k) \quad (4.8)$$

を考える。このとき、例えば任意の $\alpha \in \mathbb{R}$ に対して、 $u(k) := \alpha x(k)$ により実現される軌道 $x(k)$ は、実数列 $u(k) := \alpha(a + \alpha)^k x(0)$ により再現できる（図 4.1a）。つまり、制御入力を状態変数やその履歴に依存させる利点はない。これは制御対象や制御則が時刻に依存したり非線形であっても同様であり、確定システムの制御問題は本質的にフィードフォワード制御で十分である。 \diamond

図 4.1: (4.9) ($a = 1.1$) のフィードフォワード制御 $u_k = -0.2 \cdot 0.9^{k-1}$

つぎに確率システムの例として,

$$x_{k+1} = ax_k + u_k + 0.1v_k, \quad a > 1, \quad x_0 = 1 \quad (4.9)$$

を考える。ただし、 $\{v_k\}_{k \in \mathbb{Z}_+}$ は独立同分布で $v_k \sim \mathcal{N}(0, 1)$ とする。このとき、 $|a + \alpha| < 1$ を満たす α を用いた $u_k = \alpha x_k$ により $\text{Var}[x_k]$ は有限にとどまる（定理 3.2.4）。一方で、どのような確定的な制御入力 $u_{0:} = u_{0:}$ も x_k の分散には影響せず、 $\text{Var}[x_k] \rightarrow \infty$ である（図 4.1b）。このように、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が存在する。これは、現代制御論において確定システムに対する最適レギュレータ問題がフィードフォワード制御の形で定式化され、結果的に最適制御則が状態フィードバックの形をもつことは本質的に異なる。確率制御問題の解法において、しばしば状態非依存のフィードフォワード制御入力に限定することで計算負荷の軽減が図られるが、この差には注意する必要がある。また、入力にあえて確率性をもたせることもあるが、馴染みがないかもしれない一つ実用例をみておこう。

例 4.1.2 (ディザ量子化器)。確定システム

4.1 動的計画法 47

$$x_{k+1} = Ax_k + Bu_k, x_0 = \begin{bmatrix} 1 & 0.5 \end{bmatrix}^\top, y_k = Cx_k \quad (4.10)$$

$$A_c := \begin{bmatrix} 0 & 4 \\ -3 & 2 \end{bmatrix}, B_c := \begin{bmatrix} 0 \\ 1 \end{bmatrix}, C_c := -\begin{bmatrix} 0.3 & 4 \end{bmatrix} \quad (4.11)$$

$$A := e^{A_c T_c}, B := \int_0^{T_c} e^{A_c t} B_c dt, C := C_c, T_c = 0.01 \quad (4.12)$$

を考える。これは連続時間システム

$$\dot{x}_c = A_c x_c + B_c u_c, \quad y_c = C_c x_c \quad (4.13)$$

をサンプル時間 T_c で離散化した確定システムであり、図 4.2 に三種類の制御則に対する時間応答を示す。まず、確定状態フィードバック $u_k = y_k$ では、閉ループ系の安定性 ($\rho(A + BC) < 1$) より u, y ともに速やかに 0 に収束している。つぎに、このシステムには整数値しか入力できない状況を考える。素朴には、 \mathcal{Q} を実数入力値に対して最も近い整数を出力する一様量子化器とし、確定的状態フィードバック $u_k = \mathcal{Q}(y_k)$ を用いることが考えられるが、量子化誤差により大きな振動を生じる。つぎに、独立同分布の $z_k \sim \text{Uni}([-0.5, 0.5])$ を用いて確率的状態フィードバック $u_k = \mathcal{Q}(y_k + z_k)$ とすると、この望ましくない振動が大幅に抑えられる。このような確率性雑音を印加することで量子化誤差の影響を抑える手法はランダムディザと呼ばれる。解析方法としては、量子化誤差 $e_k := \mathcal{Q}(y_k + z_k) - y_k$ は時間的に無相関、 $\mathbb{E}[e_k] = 0, \text{Var}[e_k] \leq 0.5^2$ であることを示すことができるため、

$$x_{k+1} = (A + BC)x_k + Be_k \quad (4.14)$$

に定理 3.2.1 を適用して評価できる [11]. \diamond

さて、遷移密度関数 Ψ_k をもつ確率システムに対して、条件付き確率の定義と周辺化から

$$\varphi_{x_{k+1}}(x') = \int_{\mathbb{U}} \int_{\mathbb{X}} \Psi_k(x'|x, u) \varphi_{(x_k, u_k)}(x, u) dx du \quad (4.15)$$

が成り立つ。ここで、状態フィードバック制御則の具体的な表現に対して、

48 4. 確率制御 1 (18)

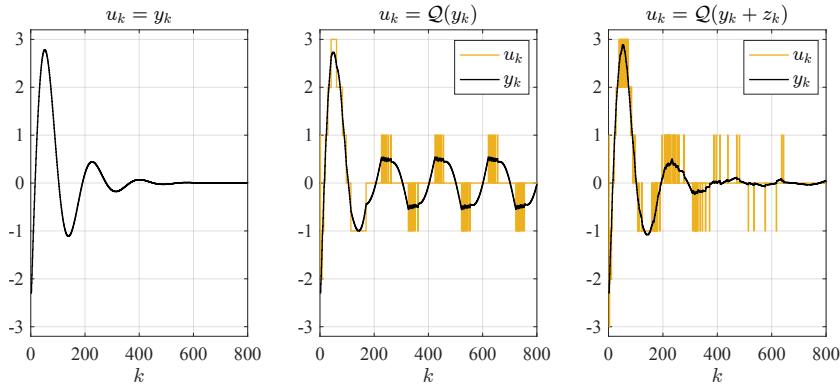


図 4.2: ランダムディザ量化

- $x_{0:k}$ の実現値から u_k の確率密度関数が定まる確率的状態履歴フィードバックが

$$\varphi_{u_k}(u|x_{0:k} = x_{0:k}) = \pi_k(u|x_{0:k}), \quad u_k \perp\!\!\!\perp_{x_{0:k}} u_{0:k-1} \quad (4.16)$$

で与えられるとき、状態-入力過程の履歴の確率密度関数は、

$$\varphi_{(x_{0:k}, u_{0:k-1})}(x_{0:k}, u_{0:k-1}) = \left\{ \prod_{i=0}^{k-1} \Psi_i(x_{i+1}|x_i, u_i) \pi_i(u_i|x_{0:i}) \right\} \varphi_{x_0}(x_0) \quad (4.17)$$

- x_k の実現値から u_k の確率密度関数が定まる確率的状態フィードバックが

$$\varphi_{u_k}(u|x_k = x) = \pi_k(u|x), \quad u_k \perp\!\!\!\perp_{x_k} (x_{0:k-1}, u_{0:k-1}) \quad (4.18)$$

で与えられるとき、対応する状態過程は遷移確率密度関数が

$$\varphi_{x_{k+1}}(x'|x_k = x) = \int_{\mathbb{U}} \Psi_k(x'|x, u) \pi_k(u|x) du \quad (4.19)$$

で与えられるマルコフ過程、

- $u_k \in \sigma(x_k)$ である確定的状態フィードバックが

$$u_k := \pi_k(x_k) \quad (4.20)$$

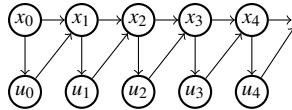


図 4.3: 状態フィードバック制御則 (4.19)

で与えられる場合の遷移確率密度関数は

$$\varphi_{x_{k+1}}(x' | x_k = x) = \Psi_k(x' | x, \pi_k(x)) \quad (4.21)$$

となる。

記法 4.1.3. これらの制御則 π により制御されたシステム（閉ループ系と呼ぶ）であることを明示する場合には、状態過程を x^π 、入力過程を u^π 、期待値を $\mathbb{E}^\pi[\cdot]$ のように表記する。また、必要に応じて $\mathbb{E}[\cdot | x_k = x]$ を $\mathbb{E}_{x_k=x}[\cdot]$ のようにあらわす。

△

つぎに、望ましい挙動を規定するために、各時刻 $k \in \mathbb{Z}_+$ および確率変数 (x_k, u_k) に対して実数値のコスト ℓ を割り当てる。この各時刻のコストの総和 J は実数値確率変数であり、その期待値を小さくすることを制御目標とする。

問題 4.1.4 (確率システムの最適制御問題 (有限時間区間)). 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、終端時刻 \bar{k} 、ステージコスト $\ell: \llbracket \bar{k} \rrbracket \times \mathbb{X} \times \mathbb{U} \rightarrow \mathbb{R}$ 、終端コスト $\ell_f: \mathbb{X} \rightarrow \mathbb{R}$ 、許容入力集合 $U_k(x) \subset \mathbb{U}$, $k \in \llbracket \bar{k} \rrbracket$, $x \in \mathbb{X}$ に対して、

$$u_k \in U_k(x_k), \forall k \in \llbracket \bar{k} \rrbracket \quad (4.22)$$

を満たし、

$$J(u_\cdot) := \mathbb{E} \left[\sum_{k=0}^{\bar{k}-1} \ell(k, x_k^u, u_k) + \ell_f(x_{\bar{k}}^u) \right] \quad (4.23)$$

を最小化する因果的な制御入力 u_\cdot を求めよ。[†]

△

[†] 必要に応じて評価関数を $\sum_{k=0}^{\bar{k}} \ell(k, x_k, u_k)$ と簡略表記するため、 $\ell(\bar{k}, x, u) := \ell_f(x)$ とする。

50 4. 確率制御 1 (18)

ここで、各時刻において状態に依存して入力できる値が限定される状況を表現するために、集合値関数 U_k を導入した。たとえば、入力値として整数しか許していない例 4.1.2 の場合、 $U_k(x) = \mathbb{Z}$ とかける。以下では、入力過程 u_k が因果性および制約 (4.22) をすべての時刻で満足していることを強調する場合は、

$$u_k \in \mathcal{U}_a \quad (4.24)$$

などと略記する。問題 4.1.4 において制御対象とコストはともに過去の履歴に依存しないため、制御入力を状態や入力の過去の履歴にもとづいて決定する必要はないと考えられるが、これを確認しておく。

定理 4.1.5 (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 、初期状態 $x_0 \in \text{rv}(\mathbb{X})$ をもつ確率システムにおいて、任意の因果的な入力 u_k^h に対して、対応する状態過程を x_k^h とおく。また、 $\varphi_{x_k^h}(x) > 0$ を満たす (k, x) に対して、

$$\pi_k^M(u|x) := \varphi_{u_k^h}(u|x_k^h = x), \quad \forall (k, x, u) \in \mathbb{Z}_+ \times \mathbb{X} \times \mathbb{U} \quad (4.25)$$

による確率的状態フィードバックの状態・入力過程を x_k^M, u_k^M とおく。このとき、任意の $k \in \mathbb{Z}_+$ に対して (x_k^h, u_k^h) と (x_k^M, u_k^M) は同分布をもつ。 \triangleleft

証明。 まず、初期状態 x_0 は共通であるため、 $x_0^h = x_0^M$ である。一方で、 x_k^h と x_k^M が同分布をもつとき、

$$\varphi_{(x_k^h, u_k^h)}(x, u) = \varphi_{u_k^h}(u|x_k^h = x)\varphi_{x_k^h}(x) = \pi_k^M(u|x)\varphi_{x_k^M}(x) = \varphi_{(x_k^M, u_k^M)}(x, u)$$

から、 (x_k^h, u_k^h) と (x_k^M, u_k^M) は同分布をもつ。さらに、(4.15) は制御則に依存せず共通であることから、 x_{k+1}^h と x_{k+1}^M が同分布をもつ。これらを帰納的に用いることで、求める結果がしたがう。 \square

この定理は、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であることを意味する。特に、問題 4.1.4 もこれに該当する ($\mathbb{E}[\ell(k, x_k, u_k)] = \iint \ell(k, x, u) \varphi_{(x_k, u_k)}(x, u) dudx$)。

ただし、最適状態フィードバック制御則が初期状態 x_0 の分布に依存するかについては、ここでは言及していない（定理 4.2.1）。また、任意時刻での (x_k^h, u_k^h) と (x_k^M, u_k^M) の同分布性は経路 $(x_{0:}^h, u_{0:}^h)$ と $(x_{0:}^M, u_{0:}^M)$ の同分布性を意味しない（注意 3.1.9 参照）ことに注意する。

演習 4.1. (4.23) ではあらわせない評価関数

$$J_1(u_{\cdot}) := \mathbb{E} \left[\sum_{k=1}^{\bar{k}} x_{k-1}^u x_k^u \right], \quad J_2(u_{\cdot}) := \mathbb{E} \left[\max_{k \in [\bar{k}]} \|x_k^u\| \right], \quad J_3(u_{\cdot}) := \sum_{k=0}^{\bar{k}} H(x_k^u)$$

を最小化する制御問題に対して、履歴非依存の確率的状態フィードバック則で十分か判定せよ。ただし、 $H(\cdot)$ は定義 2.2.2 のエントロピーである。 \triangleleft

確定的なシステムに対する非線形最適制御問題の代表的解法に、動的計画法によりベルマン方程式を導く手法と、ポントリヤーギンの最大値原理によりオイラー・ラグランジュ方程式を導く手法がある [12]。本書では、強化学習でも中心的な役割を果たす動的計画法に関する結果を詳細に導入し、後者に関しては 4.2.3 節にて簡単に紹介することにとどめる。なお、どちらにおいても導出の背景にある考え方や最終的に得られる条件の形は、確定システムの場合と大きな差異はない。

4.1.2 ベルマン方程式

この節では、問題 4.1.4 における初期時刻 $k = 0$ を一般化し、初期状態が一点で与えられる問題を考える。すなわち、**状態価値関数**（state value function）と呼ばれる

$$V^*(k, x) := \inf_{u_{k:} \in \mathcal{U}_a} \underbrace{\mathbb{E}_{x_k=x}^u \left[\sum_{i=k}^{\bar{k}-1} \ell(i, x_i, u_i) + \ell_f(x_{\bar{k}}) \right]}_{=: V^u(k, x)}, \quad k \in \mathbb{Z}_+, x \in \mathbb{X} \quad (4.26)$$

を定義する。制御対象が確定システム $x(k+1) = f(x(k), u(k))$ であるとき、最適化の時間的入れ子構造に注目する**動的計画法**（dynamic programming）の基本的な考え方によると、**ベルマン方程式**（Bellman equation）と呼ばれる

$$V^*(k, x) = \inf_{u \in U_k(x)} (\ell(k, x, u) + V^*(k+1, f(x, u))) \quad (4.27)$$

が成立していた。以下では、任意の k, x に対して $V^*(k, x)$ は有限の値をとると仮定し、状態遷移と制御則の確率性と制御則の因果性を丁寧に確認しつつ、確率システムに対する同様の結果を導出する。

まず、 V^* は本来の初期状態 x_0 の分布に依存しない関数であるが、つぎの結果により $J(u_::)$ と関連付けられる。

補題 4.1.6. 状態価値関数は

$$\inf_{u_:: \in \mathcal{U}_a} J(u_::) = \mathbb{E}[V^*(0, x_0)] \quad (4.28)$$

を満たす。 \triangleleft

証明. 定理 2.3.3 の塔特性より、

$$\begin{aligned} \inf_{u_:: \in \mathcal{U}_a} J(u_::) &= \inf_{u_:: \in \mathcal{U}_a} \int_{\mathbb{X}} V^u(k, x) \varphi_{x_0}(x) dx \\ &\geq \int_{\mathbb{X}} \inf_{u_:: \in \mathcal{U}_a} V^u(k, x) \varphi_{x_0}(x) dx = \mathbb{E}[V^*(0, x_0)] \end{aligned}$$

が成り立つ。一方で、任意の $x \in \mathbb{X}$ よび $\varepsilon > 0$ に対して、 $V^{u^x}(0, x) < V^*(0, x) + \varepsilon$ を満たす $u_::^x$ がとれる。ここで、 x_0 の実現値 x_0 に依存して $u_::^{x_0}$ を用いる \bar{u} ($\bar{u}_::|_{x_0=x_0} = u_::^{x_0}$) は因果性に反しない。また、 $J(\bar{u}_::) = \int_{\mathbb{X}} V^{\bar{u}^x}(0, x) \varphi_{x_0}(x) dx \leq \mathbb{E}[V^*(0, x_0)] + \varepsilon$ であり、 ε の任意性より求める結果がしたがう。 \square

つぎに、(4.27) に現れる $V(k+1, f(x, u))$ は、確率システムにおいては

$$\begin{aligned} \mathcal{A}_x[V](k, x, u) &:= \mathbb{E}[V(k+1, x_{k+1}) | x_k = x, u_k = u] \quad (4.29) \\ &= \int_{\mathbb{X}} V(k+1, x') \Psi_k(x' | x, u) dx' \end{aligned}$$

が対応する。これらの準備のもと、状態価値関数に含まれる最小化の階層的な構造より、最適制御において中心的な役割を果たすつぎの定理が得られる。

定理 4.1.7 (ベルマン方程式). 問題 4.1.4 に対して

$$V(\bar{k}, \mathbf{x}) = \ell_f(\mathbf{x}), \forall \mathbf{x} \in \mathbb{X}, \quad (4.30a)$$

$$V(k, \mathbf{x}) = \inf_{\mathbf{u} \in U_k(\mathbf{x})} (\ell + \mathcal{A}_x[V])(k, \mathbf{x}, \mathbf{u}), \forall (k, \mathbf{x}) \in [\![\bar{k}]\!] \times \mathbb{X} \quad (4.30b)$$

の解 V は (4.26) で定義される状態価値関数 V^* に一致する. \triangleleft

証明. まず、(4.26) の右辺において、制御入力の決定およびステージコストを分割し、初期時刻とそれ以降の最適化を分離して評価する。具体的には、定理 2.3.3 の塔特性を用いることで、

$$\begin{aligned} V^*(k, \mathbf{x}) &= \inf_{u_k} \inf_{u_{k+1:}} \mathbb{E}_{x_k=\mathbf{x}} \left[\ell(k, \mathbf{x}, u_k) + \sum_{i=k+1}^{\bar{k}} \ell(i, x_i^u, u_i) \right] \\ &= \inf_{u_k} \inf_{u_{k+1:}} \int_{U_k(\mathbf{x})} \left\{ \ell(k, \mathbf{x}, \mathbf{u}) + \mathbb{E}_{x_k=\mathbf{x}, u_k=\mathbf{u}} \left[\sum_{i=k+1}^{\bar{k}} \ell(i, x_i^u, u_i) \right] \right\} \varphi_{u_k|x_k=\mathbf{x}}(\mathbf{u}) d\mathbf{u} \\ &= \inf_{u_k} \int_{U_k(\mathbf{x})} \left\{ \ell(k, \mathbf{x}, \mathbf{u}) + \inf_{u_{k+1:}} \mathbb{E}_{x_k=\mathbf{x}, u_k=\mathbf{u}} \left[\sum_{i=k+1}^{\bar{k}} \ell(i, x_i^u, u_i) \right] \right\} \varphi_{u_k|x_k=\mathbf{x}}(\mathbf{u}) d\mathbf{u} \end{aligned}$$

が成り立つ。ただし、補題 4.1.6 と同様に、 $u_{k+1:}$ が u_k の実現値に依存することは因果性に反さないことから、下界と積分の順序を交換した。さらに、塔特性および x_i のマルコフ性より

$$\begin{aligned} &\inf_{u_{k+1:}} \mathbb{E}_{x_k=\mathbf{x}, u_k=\mathbf{u}} \left[\sum_{i=k+1}^{\bar{k}} \ell(i, x_i^u, u_i) \right] \\ &= \inf_{u_{k+1:}} \int_{\mathbb{X}} \mathbb{E}_{x_{k+1}=x'}^u \left[\sum_{i=k+1}^{\bar{k}} \ell(i, x_i, u_i) \right] \Psi_k(x'|x, u) dx' \\ &= \int_{\mathbb{X}} \underbrace{\inf_{u_{k+1:}} \mathbb{E}_{x_{k+1}=x'}^u \left[\sum_{i=k+1}^{\bar{k}} \ell(i, x_i, u_i) \right]}_{V^*(k+1, x')} \underbrace{\Psi_k(x'|x, u)}_{u_{k+1:} \text{に依存しない}} dx' \\ &= \mathcal{A}_x[V^*](k, \mathbf{x}, \mathbf{u}) \end{aligned}$$

が成り立つ。ここでも、 $u_{k+1:}$ が x_{k+1} の実現値に依存することは因果性に反さないことから、下界と積分の順序を交換した。したがって、

$$V^*(k, x) = \inf_{u \in \text{rv}(U_k(x))} \mathbb{E} [(\ell + \mathcal{A}_x[V])(k, x, u)] \quad (4.31)$$

が得られる（ここでの u は $u_k|_{x_k=x}$ に対応する）。

最後に、任意の $u \in \text{rv}(U_k(x))$ および V に対して

$$\int_{U_k(x)} (\ell + \mathcal{A}_x[V])(k, x, u) \varphi_u(u) du \geq \inf_{u \in U_k(x)} (\ell + \mathcal{A}_x[V])(k, x, u) \quad (4.32)$$

が成り立つことに注意すると、(4.31) の右辺は確定的な入力値 $u \in U_k(x)$ に限定した最小化に置き換えられる。□

ベルマン方程式 (4.30b) は関数の漸化式であり、終端コストから時間逆向きに状態価値関数を求められることが定理 4.1.7 の主張である。こうしたベルマン方程式や最適制御則の性質を 4.2 節にて議論する。

4.1.3 【オプション】リスク鋭敏型制御

本節では、問題 4.1.4 の枠組みには当てはまらない問題のうち、ロバスト制御や統計力学的な視点においても頻繁に登場するつぎの問題に関する結果を紹介する。

問題 4.1.8 (リスク鋭敏型制御). 問題 4.1.4 の確率システム、状態依存入力制約のもと、 $\tau > 0$ に対する評価関数

$$J(u_\cdot) := \frac{1}{\tau} \log \mathbb{E} \left[\exp \left\{ \tau \left(\sum_{k=0}^{\bar{k}-1} \ell(k, x_k, u_k) + \ell_f(x_{\bar{k}}) \right) \right\} \right] \quad (4.33)$$

を最小化する因果的な制御入力 u_\cdot を求めよ。△

ここで対数をとらなくても問題としては等価であるが、期待値と指數関数の順序には注意する必要がある[†]。この問題設定は、 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ため、リスク鋭敏型制御 (risk sensitive control) と呼ばれる。また、 $(1/\tau) \log \mathbb{E} [\exp \tau \Phi] = \mathbb{E} [\Phi] + \tau \text{Var} [\Phi] + O(\tau^2)$ より、 $\tau \rightarrow 0$ の

[†] 以降、`log_sum_exp` の形の関数が多数現れる。なお、 $f(x) = \log(\sum_i e^{x_i})$ は凸関数である。

極限で問題 4.1.4 に一致し、コストの期待値だけでなくその分散も小さくすることを目指している。ここで、ステージコストの和の最小化として定式化されていた問題 4.1.4 と比較して、(4.33) のコスト関数は $\exp(\tau\ell(k, x_k, u_k))$ の積で与えられている。しかし、コスト関数が時刻ごとに分離できるという性質は共通であることに注意すると、定理 4.1.7 と同様の議論により、問題 4.1.8 に対するベルマン方程式が得られる。

定理 4.1.9. 問題 4.1.8 に対して、

$$V(\bar{k}, x) = \exp(\tau\ell_f(x)), \quad \forall x \in \mathbb{X}, \quad (4.34a)$$

$$V(k, x) = \inf_{u \in U_k(x)} \{\exp(\tau\ell(k, x, u)) \mathcal{A}_x[V](k, x, u)\}, \quad \forall k \in \llbracket \bar{k} \rrbracket, \quad x \in \mathbb{X} \quad (4.34b)$$

の解 V は (4.26) と同様に定義される状態価値関数 V^* に一致する。 ◇

証明. (4.33) の期待値の中の関数について、

$$\exp\left(\tau \sum_{j=k}^{\bar{k}} \ell(j, x_j, u_j)\right) = \exp(\tau\ell(k, x_k, u_k)) \exp\left(\tau \sum_{j=k+1}^{\bar{k}} \ell(j, x_j, u_j)\right)$$

に注意すると、定理 4.1.7 の導出と同じ議論により、求める結果がしたがう。 □

問題 4.1.4 と同様に、リスク鋭敏型制御問題も、対象を線形システムに限定すると最適解が簡明な表現をもつ [13]。

4.2 最適制御則

本節では、4.2.1 節において状態価値関数の上下限を与える関数不等式を用いて状態価値関数や最適制御則の性質を整理したのちに、4.2.2 節において線形システムに対する結果を紹介する。

4.2.1 ベルマン方程式の不等式評価

状態価値関数 V^* が与えられているとき、定理 4.1.7 の証明より、各 x に対し

56 4. 確率制御 1 (18)

て (4.30b) の最小化を達成する u が、最適な確定的状態フィードバック入力を与える。つぎの結果は状態価値関数の上限[†]を用いて、制御則が具体的に満たすべき条件を与える。

定理 4.2.1 (最適制御則). 問題 4.1.4において、 $\hat{V} : [\bar{k}] \times \mathbb{X} \rightarrow \mathbb{R}$, $\pi_k : \mathbb{X} \rightarrow \mathbb{U}$, $\varepsilon \geq 0$ が、

$$\hat{V}(\bar{k}, x) \geq \ell_f(x), \quad \forall x \in \mathbb{X}, \quad (4.35a)$$

$$\hat{V}(k, x) + \varepsilon \geq (\ell + \mathcal{A}_x[\hat{V}](k, x, \pi_k(x))), \quad \pi_k(x) \in U_k(x), \quad \forall k \in [\bar{k}], \forall x \in \mathbb{X}, \quad (4.35b)$$

を満たすとする。このとき、確定的状態フィードバック $u_k = \pi_k(x_k)$ は

$$\mathbb{E}^\pi \left[\sum_{k=0}^{\bar{k}} \ell(k, x_k, u_k) \right] \leq \mathbb{E}_{x_0}[\hat{V}(0, x_0)] + \varepsilon \bar{k} \quad (4.36)$$

を満たす。また、 \hat{V} が状態価値関数 V^* のとき、任意の $\varepsilon > 0$ に対して条件を満たす $\pi_{0:\bar{k}-1} := (\pi_0, \dots, \pi_{\bar{k}-1})$ が存在する。 ◇

証明. まず、

$$V^\pi(k, x) := \mathbb{E}_{x_k=x}^\pi \left[\sum_{i=k}^{\bar{k}} \ell(i, x_i, u_i) \right] \quad (4.37)$$

の定義のもと、(4.36) の左辺は $\mathbb{E}_{x_0}[V^\pi(0, x_0)]$ に等しい。したがって、

$$V^\pi(k, x) \leq \hat{V}(k, x) + \varepsilon(\bar{k} - k), \quad \forall k \in [\bar{k}], \quad \forall x \in \mathbb{X} \quad (4.38)$$

を示せば十分である。まず $k = \bar{k}$ のとき、 $V^\pi = \ell_f$ より自明に成立する。また、ある k について成立すると仮定すると、

$$\begin{aligned} V^\pi(k-1, x) &= \ell(k-1, x, \pi_{k-1}(x)) + \mathcal{A}_x[V^\pi](k-1, x, \pi_{k-1}(x)) \\ &\leq \ell(k-1, x, \pi_{k-1}(x)) + \mathcal{A}_x[\hat{V}](k-1, x, \pi_{k-1}(x)) + \varepsilon(\bar{k} - k) \\ &\leq \hat{V}(k-1, x) + \varepsilon(\bar{k} - k + 1) \quad (\because (4.35b)) \end{aligned}$$

[†] 本書を通じて、 \hat{V} は状態価値関数 V^* の上限 (\check{V} は下限) に関する関数をあらわす。

が成り立つので、再帰的に $k \in \llbracket \bar{k} \rrbracket$ に対して (4.38) がしたがう。最後の主張は、 V^* がベルマン方程式 (4.30) を満たすことと下界の定義から明らか。□

問題 4.1.4 は与えられた初期状態分布に対する挙動を制御する問題として定式化されているが、この定理は「任意の初期状態分布の制御問題に対して準最適な制御則」が構成できることを意味する。また、問題 4.1.4 に対する制御則を設計するためだけであれば、制御則に確率性を導入する必要はない[†]。

ところで、定理 4.2.1 の 4.35 には状態価値関数 V^* はあらわれていないにもかかわらず、 \check{V} は状態価値関数の上界の評価

$$V^*(k, x) \leq \check{V}(k, x) + \varepsilon(\bar{k} - k) \quad (4.39)$$

を与える。逆に、下界に関する不等式を与えて本節を締めくくる。

系 4.2.2 (状態価値関数の下限)。問題 4.1.4において、関数 $\check{V} : \llbracket \bar{k} \rrbracket \times \mathbb{X} \rightarrow \mathbb{R}_+$ が

$$\check{V}(\bar{k}, x) \leq \ell_f(x), \forall x \in \mathbb{X}, \quad (4.40a)$$

$$\check{V}(k, x) \leq (\ell + \mathcal{A}_x[\check{V}])(k, x, u), \forall k \in \llbracket \bar{k} \rrbracket, x \in \mathbb{X}, u \in U_k(x), \quad (4.40b)$$

を満たすとする。このとき、(4.26) の状態価値関数 V^* に対して

$$\check{V}(k, x) \leq V^*(k, x), \forall k \in \llbracket \bar{k} \rrbracket, x \in \mathbb{X} \quad (4.41)$$

である。△

証明。 $k = \bar{k}$ の場合は、(4.40a) そのものである。また、ある $k \geq 1$ に対して (4.41) が成り立つと仮定すると、

$$\begin{aligned} \check{V}(k-1, x) &\leq \inf_{u \in U_{k-1}(x)} (\ell + \mathcal{A}_x[\check{V}])(k-1, x, u) \\ &\leq \inf_{u \in U_{k-1}(x)} (\ell + \mathcal{A}_x[V^*])(k-1, x, u) = V^*(k-1, x) \end{aligned}$$

が成り立つ。ただし、最後の等式は定理 4.1.7 のベルマン方程式である。□

[†] 例 4.1.2 のランダムディザでは確率性を有効活用しており、10.1.2 節の制御問題の最適制御則も確率的である。

4.2.2 線形二次レギュレータ

\mathbb{X} が \mathbb{R}^{n_x} のように連続値をとる集合である場合、ベルマン方程式は無限個の等式条件であり直接的に解くことは困難である^{†1}。そこで、実数値集合を十分大きな N 個の点で近似する^{†2}ことも考えられるが、 \mathbb{R}^{n_x} の近似には N^{n_x} 個の点が必要となり、 n_x が大きいとき現実的な計算量や記憶容量で最適化などはおこなえない。このように、状態の次元に対して必要な記憶容量や計算量が指数的に増大する問題を一般に次元の呪いと呼ぶ。また、(近似) 双模倣性 (Bi-Simulation) など連続値をダイナミクスを有限状態のダイナミクスで近似する手法も盛んに研究されているが、一般に難しい問題である [14]。本節では、状態空間の次元が大きくても効率的にベルマン方程式の解が得られる数少ない例である線形二次レギュレータ問題を考察する。

問題 4.2.3 (有限区間線形二次レギュレータ問題)。線形確率システム

$$x_{k+1} = A_k x_k + B_k u_k + v_k, \quad (4.42)$$

$$\mathbb{E}[v_k] = 0, \mathbb{E}[v_k v_k^\top] = N_k \quad (4.43)$$

に対して、 $U_k(x) := \mathbb{R}^{n_u}$ およびステージコストが

$$\begin{bmatrix} Q_k & S_k^\top \\ S_k & R_k \end{bmatrix} \succeq O, R_k \succ O, \Pi_{\bar{k}} \succeq O \quad (4.44)$$

を用いた二次形式

$$\ell(k, x, u) = \begin{bmatrix} x^\top & u^\top \end{bmatrix} \begin{bmatrix} Q_k & S_k^\top \\ S_k & R_k \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix}, \quad \ell_f(x) = x^\top \Pi_{\bar{k}} x \quad (4.45)$$

により与えられる問題 4.1.4 を有限区間線形二次レギュレータ (Linear Quadratic Regulator; LQR) 問題という。 ◇

^{†1} 11.1 節において、凸最適化を用いて連続状態のベルマン方程式を解く手法を紹介する。

^{†2} 10 章において、有限状態のシステムの性質を概説する。

4.2 最適制御則 59

以下に示すように、この問題に対するベルマン方程式に含まれる u に関する最小化は平方完成により自明に解くことができ、状態価値関数は二次形式で与えられる。

定理 4.2.4 (有限区間 LQR 問題の最適解)。問題 4.2.3において、

$$\tilde{R}_k := R_k + B_k^\top \Pi_k B_k, \quad \tilde{S}_k := S_k + B_k^\top \Pi_k A_k, \quad \tilde{Q}_k := Q_k + A_k^\top \Pi_k A_k \quad (4.46)$$

$$\Pi_{k-1} := \tilde{Q}_k - \tilde{S}_k^\top \tilde{R}_k^{-1} \tilde{S}_k, \quad K_k := \tilde{R}_k^{-1} \tilde{S}_k \quad (4.47)$$

$$\gamma_k := 0, \quad \gamma_k := \sum_{i=k}^{\bar{k}-1} \text{Trace}(N_i \Pi_{i+1}) \quad (4.48)$$

を定義すると、 Π_k は半正定行列であり、 x_0 に依存しない線形制御則

$$u_k := -K_k x_k, \quad (4.49)$$

は最適コスト

$$\mathbb{E}_{x_0}[x_0^\top \Pi_0 x_0] + \gamma_0 \quad (4.50)$$

を達成する。 \diamond

証明。 表記の簡単化のため、自明な下付き添え字 k は適宜省略する。まず、

$$V(k, x) := x^\top \Pi_k x + \gamma_k \quad (4.51)$$

が状態価値関数であることを示す。定理 4.1.7 より、ベルマン方程式

$$V(k-1, x) = \inf_{u \in \mathbb{R}^{n_u}} \{ \ell_{k-1}(x, u) + \mathbb{E}[V(k, Ax + Bu + v_{k-1})] \} \quad (4.52)$$

を確認すればよい。まず、 $\gamma_{k-1} = \gamma_k + \text{Trace}(N_{k-1} \Pi_k)$ より、期待値部分は

$$\begin{aligned} & \mathbb{E}[(Ax + Bu + v_{k-1})^\top \Pi_k (Ax + Bu + v_{k-1})] + \gamma_k \\ &= (Ax + Bu)^\top \Pi_k (Ax + Bu) + \mathbb{E}[v_{k-1}^\top \Pi_k v_{k-1}] + \gamma_k \\ &= (Ax + Bu)^\top \Pi_k (Ax + Bu) + \gamma_{k-1} \end{aligned}$$

(\because トレースの巡回性より $\text{Trace}(v_{k-1}^\top \Pi_k v_{k-1}) = \text{Trace}(\Pi_k v_{k-1} v_{k-1}^\top)$)

である。また、(4.47) と同様の記号のもと代入して整理すると、

$$\begin{aligned} & \ell_{k-1}(x, u) + (Ax + Bu)^\top \Pi_k (Ax + Bu) \\ &= (u + \tilde{R}^{-1} \tilde{S}x)^\top \tilde{R} (u + \tilde{R}^{-1} \tilde{S}x) + x^\top \Pi_{k-1} x \end{aligned} \quad (4.53)$$

を得る。この等式から、

- (4.52) が成り立ち、 $u = \pi_k(x) := -\tilde{R}^{-1} \tilde{S}x$ が最小値を与える
- ℓ_{k-1} が非負かつ $\Pi_k \succeq O$ ならば $\Pi_{k-1} \succeq O$ である

が示される。最後に、 $\hat{V} := V^*$ および $\varepsilon = 0$ は定理 4.2.1 の条件を満たすため、求める結果がしたがう。 \square

この定理において、雑音 v_k により最適制御性能は劣化する ($\gamma_0 \geq 0$) が、最適制御則には影響しない。これは**確実性等価原理** (certainty-equivalent principle) と呼ばれ、線形二次レギュレータ特有の性質である。

注意 4.2.5 (モデル予測制御)。各時刻 k において、区間長 \bar{k} の有限時間最適制御問題を実時間で解くことで最適制御入力 $u_{i|k}^*$, $i = k, k+1, \dots, k+\bar{k}$ を求め、その初期時刻での値 $u_{k|k}^*$ を現時刻での制御入力として用いる手法を**モデル予測制御** (model predictive control) と呼ぶ [15]。制御対象が線形システムであっても、入力や状態の取りうる値に制約がある場合、定理 4.2.4 のように簡単には最適制御則を与えることはできない。非線形モデル予測制御の近年の動向は [16] によくまとめられている。微分動的計画法や iLQR など、LQR の結果を活用して非線形制御問題を解く手法も広く用いられる [17, 18]。 \triangleleft

4.2.3 【オプション】最適性原理

確定離散時間システムに対する未定乗数法にもとづく解法を紹介し、確率システムへの拡張はその困難性の要因の記述に留める。確定システム

$$x_{k+1} = f_k(x_k, u_k), k \in \llbracket \bar{k} \rrbracket \quad (4.54)$$

においてに対して問題 4.1.4 を考える。これは等式拘束条件 (4.54) を満たす決

定変数 x_{\cdot}, u_{\cdot} をもつ最適化問題にすぎない。そこで、ラグランジュの未定乗数 $\zeta_k \in \mathbb{R}^{n_x}$ を導入し、

$$L(x_{1:\bar{k}}, u_{0:\bar{k}-1}, \zeta_{0:\bar{k}-1}) := \ell_f(x_{\bar{k}}) + \sum_{k=0}^{\bar{k}-1} \left\{ \ell(k, x_k, u_k) - \zeta_k^\top (x_{k+1} - f_k(x_k, u_k)) \right\} \quad (4.55)$$

を定義する。ここで (4.54) を満たす (x_{\cdot}, u_{\cdot}) が L の停留点、すなわち

$$\sum_{k=1}^{\bar{k}} \frac{\partial L}{\partial x_k}(x_{1:\bar{k}}, u_{0:\bar{k}-1}, \zeta_{0:\bar{k}-1}) \delta_{x_k} + \sum_{k=0}^{\bar{k}-1} \frac{\partial L}{\partial u_k}(x_{1:\bar{k}}, u_{0:\bar{k}-1}, \zeta_{0:\bar{k}-1}) \delta_{u_k} = 0 \quad (4.56)$$

が任意の微小変動 $\delta_{x_k} \in \mathbb{R}^{n_x}$, $\delta_{u_k} \in \mathbb{R}^{n_u}$ に対して成り立つ条件を書き下すと、

$$\zeta_{k-1} = \frac{\partial f_k}{\partial x}(x_k, u_k)^\top \zeta_k + \frac{\partial \ell}{\partial x}(k, x_k, u_k), \quad k = 1, 2, \dots, \bar{k}-1 \quad (4.57)$$

$$\zeta_{\bar{k}-1} = \frac{\partial \ell_f}{\partial x}(x_{\bar{k}})^\top, \quad (4.58)$$

$$\frac{\partial \ell}{\partial u}(k, x_k, u_k) + \zeta_k^\top \frac{\partial f_k}{\partial u}(x_k, u_k) = 0, \quad k \in [\bar{k}] \quad (4.59)$$

が得られる。したがって、これらの方程式を満たす ζ_{\cdot} が存在することが (x_{\cdot}, u_{\cdot}) が停留点となるための必要条件である。一般に ζ_{\cdot} を随伴変数 (adjoint variable) もしくは共状態 (co-state)，逆時間で発展する (4.57) を随伴システム (adjoint system) と呼ぶ。もしも，(4.59) を u_k について解くことができれば，これを用いて (4.54) および (4.57) から u_k を消去して， x_{\cdot}, ζ_{\cdot} に関する二点境界値問題 ($x_0 = x_0$ および (4.58)) に帰着する。

ここで， $\delta_{u_k} \in \mathbb{R}^{n_u}$ は実ベクトルを考えれば十分であったのに対して，確率システム制御の場合には u_k は確率変数であり，その微小変動 δ_{u_k} も u_k と同じ因果性を有する確率変数を考える必要が生じ，(4.56) は期待値の意味で 0 であればよい。また，随伴変数も同様に確率変数を考える必要があり，時間発展が時間逆向きであることと，因果性との整合性にも注意を払う必要がある。その点にさえ注意すれば，確率システムの場合にも同様の関係式を導出することは可能であるが，本書では割愛するため文献 [19] などを参照してほしい。

5 状態推定 (17)

実世界の現象をモデリングする際、状態変数に対応する物理量をすべて測定できることは稀である。したがって、これらの物理量のうち直接測定できない内部情報を測らずして知ることができれば、疑似的なセンサ（ソフトセンサと呼ばれる）の役割を果たし、大きな工学的意義がある。本章では、この状態推定問題に関して、観測と制御の双対性、最適状態推定、出力フィードバック制御などの結果を紹介する。

5.1 状態推定

5.1.1 統計的推論としての定式化

現代制御論の状態方程式における出力方程式にあたる式を導入した

$$\begin{cases} x_{k+1} = f_k(x_k, u_k, v_k), & k \geq 0, \\ y_k = h_k(x_k, w_k), \end{cases} \quad (5.1)$$

を入力 u 、状態 x 、出力 y をもつ部分観測 (partially observable) 確率システムと呼ぶ。ただし、観測雑音 $w_k \in \text{rv}(\mathbb{W})$ と $h_k : \mathbb{X} \times \mathbb{W} \rightarrow \mathbb{Y}$ により定まる $y_k \in \text{rv}(\mathbb{Y})$ は、状態 x_k の部分的な情報を有する。本章では、入出力のみが実際に観測できる状況を想定し、その実現値にもとづいて状態の推論をおこなう。典型的には、制御システムに備わっているセンサー機器により取得された情報にもとづいて、直接的に測定できない物理量を推定する状況である。もしくは、近年重要なとなっているネットワークに接続された制御システムなどでは、この入出力

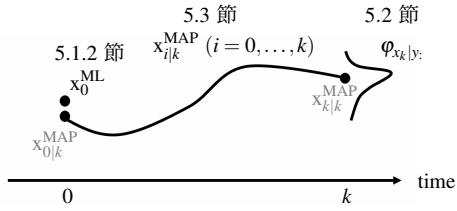


図 5.1: 状態推定の問題設定

は盗聴者が不正に取得する情報かもしれない。このとき、セキュリティのため情報を適切に秘匿するには、推定困難である方が望ましい [20].

一般に、因果性を陽に考慮して、現在時刻 k までの入出力系列 $u_{0:k}, y_{0:k}$ から、

- 現在の状態 x_k を推定することをフィルタリング (filtering),
- 未来の状態 $x_l, l > k$ を推定することを予測 (prediction),
- 過去の状態 $x_l, l < k$ を推定することをスムージング (smoothing)

とそれぞれ呼び、実時間制御にはフィルタリングを用いるなど、用途に応じて適切な推定問題が設定される。なお、現代制御論において状態推定器はオブザーバと呼ばれていたが、上記の設定別にフィルタ (filter)，予測器 (predictor)，スムーザ (smoother) と呼ぶ。

具体的な推定法も選択の自由度がある。つぎの 5.1.2 節では初期状態を推定するスムージング問題に対する最尤推定、5.2 節では予測とフィルタリング問題に対する分布推定問題、5.3 節では過去の状態軌道を MAP 推定するスムージング問題をそれぞれ検討する。得られる結果が異なるだけでなく、可観測性との関係や制御問題との双対性など、いずれも興味深い構造を有する（図 5.1）。

5.1.2 線形システムの可観測性と最尤初期状態推定

まず、現代制御論における可観測性の概念を復習しておこう。線形システム

$$x(k+1) = Ax(k), y(k) = Cx(k) \quad (5.2)$$

に対して、 $y(0), \dots, y(k-1)$ の値から $x(0)$ が一意に定まるとき、このシステムは時刻 k において可観測 (observable) であるという。このとき、可観測性行列

\mathbf{O}_k を用いた

$$\begin{bmatrix} y(0) \\ y(1) \\ \vdots \\ y(k-1) \end{bmatrix} = \mathbf{O}_k x(0), \quad \mathbf{O}_k := \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{k-1} \end{bmatrix} \quad (5.3)$$

から、システムの可観測性と可観測性グラミアン

$$\mathbf{G}_k^o := \mathbf{O}_k^\top \mathbf{O}_k = \sum_{i=0}^{k-1} A^i{}^\top C^\top C A^i \quad (5.4)$$

の正則性が等価である[†]ことが確認できる。

つぎに、この性質を確率システムの状態推定の観点から再訪しよう。部分観測確率システムが

$$x_{k+1} = Ax_k, \quad y_k = Cx_k + w_k \quad (5.5)$$

により与えるとき、(5.3) と同様の縦ベクトル表記のもと、

$$y_{0:k-1} = \mathbf{O}_k x_0 + w_{0:k-1} \quad (5.6)$$

である。ここで、観測時系列 $y_{0:k-1} = y_{0:k-1}$ が与えられたとき、 x_0 の尤度を最大化する x_0 を求める最尤推定問題の解は、(5.4) の可観測性グラミアン \mathbf{G}_k^o を用いてつぎのように特徴づけられる。

定理 5.1.1 (可観測性グラミアンとフィッシャー情報行列). (5.5)において、 $x_0 \perp\!\!\!\perp w$; かつ $w_{0:k-1} \sim \mathcal{N}(\mu_w, I)$ とする。このとき、任意の $k \in \mathbb{Z}_+$ に対して、

$$\text{FIM}_{y_{0:k-1}|x_0=x_0} = \mathbf{G}_k^o, \quad \forall x_0 \in \mathbb{R}^{n_x} \quad (5.7)$$

である。また、 \mathbf{G}_k^o が正則のとき、 $y_{0:k-1}$ にもとづく x_0 の最尤推定値は

$$x_0^{\text{ML}}(y_{0:k-1}) = (\mathbf{O}_k^\top \mathbf{O}_k)^{-1} \mathbf{O}_k^\top (y_{0:k-1} - \mu_w) \quad (5.8)$$

で与えられ、十分統計量である。 \diamond

[†] ケーリー・ハミルトンの定理より $k \geq n_x$ に対して $\text{rank } \mathbf{G}_k^o = \text{rank } \mathbf{G}_{n_x}^o$ である。

証明. 定理 2.3.16 よりしたがう. \square

この定理により、出力系列から初期状態を推定する際のフィッシャー情報行列が可観測性グラミアンに一致することが示された。また、(5.8) は最小誤差分散の意味で最適な点推定値である。線形システムが可観測であるための必要十分条件は、可観測性グラミアン G_k^o が正則であることであった。これに対応して、 G_k^o が特異な行列に近づくと、推定分散の限界の下界値であるフィッシャー情報行列の逆行列 $(G_k^o)^{-1}$ がいくらでも大きくなってしまい、初期状態の推定限界が悪化することが定理 5.1.1 の主張である。なお、(5.7) が x_0 に依存しないため、どの初期状態も推定の難しさに差はない。非線形システムに対する結果や初期状態分布が与えられている場合のベイズ推定などについては [21] を参照してほしい。

5.2 逐 次 推 定

本節では、状態変数の分布推定、すなわち $y_{0:k-1} = y_{0:k-1}$ が与えられたときに、 x_k の事後確率分布を求める問題を考える。定理 5.1.1 は利用可能な出力系列 $y_{0:k-1}$ をまとめて一括処理する形となっているが、これは系列が長くなるにつれ計算量が増加し、実時間処理には向かない。一方で、現代制御論におけるオブザーバでは、

$$x(k+1) = Ax(k) + Bu(k), \quad y(k) = Cx(k) \quad (5.9)$$

に対して、 $y(0), \dots, y(k-1), u(0), \dots, u(k-1)$ を用いて、 $x(k)$ の推定値 $\hat{x}(k)$ を求める。もしも $\hat{x}(0) = x(0)$ であれば、 $\hat{x}(k+1) = A\hat{x}(k) + Bu(k)$ という更新則により、 $x(k) = \hat{x}(k), \forall k \in \mathbb{Z}_+$ が成り立つ。また、 $x(k) \neq \hat{x}(k)$ のとき、その誤差を観測出力 $y(k)$ と状態推定値から予想される出力値 $C\hat{x}(k)$ の差を通じて定量化し、

$$\hat{x}(k+1) = A\hat{x}(k) + Bu(k) + H(y(k) - C\hat{x}(k)) \quad (5.10)$$

66 5. 状態推定 (17)

のように線形に推定値を修正すればよい。このとき、 $e(k) := x(k) - \hat{x}(k)$ が $e(k+1) = (A - HC)e(k)$ を満たすことから、 $A - HC$ がシュア安定であれば、任意の $x(0), \hat{x}(0)$ に対して $\lim_{k \rightarrow \infty} \|x(k) - \hat{x}(k)\| = 0$ が成り立つ。したがって、出力系列を保持しておく必要はなく、各時刻における処理 (5.10) も軽微である。本節では、このような逐次的な処理により状態の推定をおこなう。準備として、雑音はつぎの仮定を満たすものとする。

仮定 5.2.1 (雑音の無記憶性)。 $x_0 := \tilde{v}_{-1}$, $\tilde{v}_k := (v_k, w_k)$ の表記のもと、任意の $k \in \mathbb{Z}_+$ に対して $\tilde{v}_{-1:k-1} \perp \tilde{v}_k$ が成り立つ。 \triangleleft

この仮定のもと、入力は確定的な出力履歴フィードバック $u_k = \pi_k(y_{0:k-1}) \in \sigma(y_{0:k-1})$ (ある π_k が存在して $u_k = \pi_k(y_{0:k-1})$) とするが、その制御目的は問わない。これに対して、過去の観測系列 $y_{0:k-1}$ から現在の状態 x_k を推定する、つまり条件付き確率密度関数 $\varphi_{x_k}|_{y_{0:k-1}=y_{0:k-1}}$ を求めることが目標である。本書では統計推論にもとづく最小限の理論結果しか紹介できないため、詳細な解釈については [22] などの成書を参照してほしい。

5.2.1 線形システムの逐次予測

本節では、対象とするダイナミクスは、行列 A_k, B_k, C_k を用いて線形システム

$$\begin{cases} x_{k+1} = A_k x_k + B_k u_k + v_k, & k \geq 0, \\ y_k = C_k x_k + w_k, \end{cases} \quad (5.11)$$

で与えられ、 x_0, v_k, w_k は仮定 5.2.1 および

$$x_0 \sim \mathcal{N}(\bar{x}, \bar{\Sigma}), \quad \begin{bmatrix} v_k \\ w_k \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} N_k & L_k \\ L_k^\top & M_k \end{bmatrix}\right), \quad \begin{bmatrix} N_k & L_k \\ L_k^\top & M_k \end{bmatrix} \succ O \quad (5.12)$$

を満たすと仮定する[†]。

[†] 本節で行う密度関数に基づいたカルマンフィルタの導出では $N - LM^{-1}L^\top > 0$ の仮定が必要であるが、直交射影を用いる導出法では不要である。

定理 5.2.2 (線形システムの逐次予測). 上記の状態推定問題に対して,

$$\tilde{N}_k := N_k + A_k \Sigma_k A_k^\top, \quad \tilde{M}_k := M_k + C_k \Sigma_k C_k^\top, \quad \tilde{L}_k := L_k + A_k \Sigma_k C_k^\top \quad (5.13)$$

$$\Sigma_0 := \bar{\Sigma}, \quad \Sigma_{k+1} := \tilde{N}_k - \tilde{L}_k \tilde{M}_k^{-1} \tilde{L}_k^\top, \quad H_k := \tilde{L}_k \tilde{M}_k^{-1} \quad (5.14)$$

を定義する. このとき, $u_k \in \sigma(y_{0:k-1})$ ならば

$$\hat{x}_{k+1} = A\hat{x}_k + Bu_k + H_k(y_k - C_k\hat{x}_k), \quad \hat{x}_0 = \bar{x} \quad (5.15)$$

を用いて, $y_{0:k-1}$ で条件付けされた x_k の事後確率密度関数は,

$$x_k |_{y_{0:k-1} = \textcolor{blue}{y_{0:k-1}}} \sim \mathcal{N}(\hat{x}_k, \Sigma_k) \quad (5.16)$$

で与えられる. ただし, $y_{0:k-1}$ に対応する $u_k, \hat{x}_k \in \sigma(y_{0:k-1})$ の実現値をそれぞれ u_k, \hat{x}_k とあらわした. さらに, 推定誤差

$$e_k := x_k - \hat{x}_k \quad (5.17)$$

は $y_{0:k-1}, \hat{x}_{0:k}$ のいずれとも独立である. \triangleleft

証明. ある k に対して (5.16) が成り立つと仮定すると,

$$\begin{bmatrix} Ax_k + Bu_k \\ Cx_k \end{bmatrix} \Big|_{y_{0:k-1} = \textcolor{blue}{y_{0:k-1}}} \sim \mathcal{N} \left(\begin{bmatrix} A\hat{x}_k + Bu_k \\ C\hat{x}_k \end{bmatrix}, \begin{bmatrix} A\Sigma_k A^\top & A\Sigma_k C^\top \\ C\Sigma_k A^\top & C\Sigma_k C^\top \end{bmatrix} \right) \quad (5.18)$$

である. また, $x_k, u_k \in \sigma(x_0, v_{0:k-1}, w_{0:k-1})$ は v_k, w_k と独立であることに注意すると, 定理 2.2.10 から,

$$\begin{bmatrix} x_{k+1} \\ y_k \end{bmatrix} \Big|_{y_{0:k-1} = \textcolor{blue}{y_{0:k-1}}} \sim \mathcal{N} \left(\begin{bmatrix} A\hat{x}_k + Bu_k \\ C\hat{x}_k \end{bmatrix}, \begin{bmatrix} \tilde{N}_k & \tilde{L}_k \\ \tilde{L}_k^\top & \tilde{M}_k \end{bmatrix} \right) \quad (5.19)$$

である. これに定理 2.3.4 および 2.3.5 を適用すると,

$$\varphi_{x_{k+1}} |_{y_{0:k} = \textcolor{blue}{y_{0:k}}} \sim \mathcal{N}(\hat{x}_{k+1}, \Sigma_{k+1}) \quad (5.20)$$

が得られる. 独立性の証明は, 演習 2.2 と同様である. \square

68 5. 状態推定 (17)

この定理により、線形確率システムに対しては、現代制御論のオブザーバと同じ形で状態の逐次推定をおこなえることが示された。定理 4.2.4 と定理 5.2.2 の類似性は、しばしば制御と推定の双対性と呼ばれる。なお、事後分布の分散 Σ_k は観測値には依存しない。この事実から導かれた推定誤差の独立性は、制御と推定の分離定理を示す際に重要な役割を果たす。

5.2.2 線形二次ガウシアン制御

4 章では、状態を直接観測して入力を決定できるという仮定（状態フィードバック）のもとで、最適制御則を導出した。ここでは、過去の出力時系列のみを用いて入力を決定する出力フィードバック制御を考える。現代制御論ではオブザーバ併合形と呼ばれる制御則が考えられていた。すなわち、(5.9) に対して、 $A - BK$ がシュア安定となる行列 K が得られていると仮定する。このとき、 $u(k) = -Kx(k)$ により $x(k) \rightarrow 0$ が達成されるが、出力 $y(k)$ しか観測できない場合でも (5.10) の推定値を用いて $u(k) = -K\hat{x}(k)$ を印加すれば $x(k) \rightarrow 0$ が達成されることを容易に証明できる。同様に、確率システムに対して、 $u_k \in \sigma(y_{0:k-1})$ のもとでの最適制御問題を考える。4 章において、最適な制御則は $u_k = \pi_k(x_k)$ の形の確定的状態フィードバックにより与えられることを示したので、素朴に考えると $y_{0:k-1}$ のもとでの x_k の点推定値を \hat{x}_k とおき、 $u_k = \pi_k(\hat{x}_k)$ により適切な制御性能が達成されることが期待される（逐次推定の場合のグラフィカルモデルを図 5.2 に示す）。実際、この考え方は実用上有益であるだけでなく、対象が線形の場合は理論的に最適性が示される。

問題 5.2.3 (線形二次ガウシアン (Linear Quadratic Gaussian; LQG) 制御問題).

定理 5.2.2 と同じ部分観測線形確率システムに対して、問題 4.2.3 と同じコスト関数を最小化する出力履歴フィードバック $u_k \in \sigma(y_{0:k-1})$ を求めよ。 ◀

これは問題 4.2.3 とは設定が異なるため、定理 4.1.5 と同様の履歴非依存制御則 $u_k \in \sigma(y_{k-1})$ に限定して議論をすすめることはできない。しかし、つぎに示すように、定理 5.2.2 の状態推定器を制御対象とみなして定理 4.2.4 を適用する

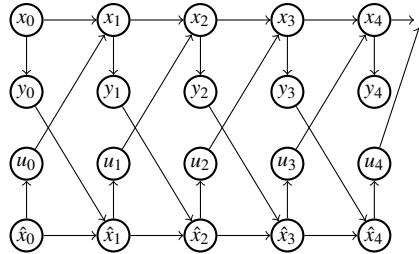


図 5.2: 状態フィードバックと逐次推定フィルタの併合系

ことができる。

定理 5.2.4 (LQG 制御問題に対する推定と制御の分離定理). 問題 5.2.3 に対して、定理 4.2.4 および定理 5.2.2 の記法のもと、最適な制御則は

$$u_k = -K(\Pi_k)\hat{x}_k \in \sigma(y_{0:k-1}) \quad (5.21)$$

で与えられ、最適コストは

$$\hat{\gamma}_k := \sum_{i=k}^{\bar{k}-1} \text{Trace}(\hat{N}_i \Pi_{i+1}) + \sum_{i=k}^{\bar{k}} \text{Trace}(Q_i \Sigma_i) \quad (5.22)$$

$$\hat{N}_k := N_k + A_k \Sigma_k A_k^\top - \Sigma_{k+1} \quad (5.23)$$

の定義のもと、定理 4.2.4 と同様に $\mathbb{E}_{x_0}[x_0^\top \Pi_0 x_0] + \hat{\gamma}_0$ で与えられる。 \diamond

証明. 条件付き期待値 \hat{x}_k の更新式 (5.15) を用いて、状態 x_k の挙動を記述する。

すなわち、(5.17) の状態推定誤差 e_k を用いると、

$$\hat{x}_{k+1} = A\hat{x}_k + Bu_k + \underbrace{H_k(Ce_k + w_k)}_{=: \hat{v}_k} \quad (5.24)$$

と書ける。定理 4.2.4 における因果性に関する仮定については (4.4) に対応する

$$(\hat{x}_{0:k}, u_{0:k}) \perp\!\!\!\perp \hat{v}_{k+1} \quad (5.25)$$

を確認すればよい。実際、(5.15) から $\hat{x}_{0:k} \in \sigma(u_{0:k-1}, y_{0:k-1})$ であることに注意すると、任意の $u_{0:k} \in \sigma(y_{0:k-1})$ に対して $(\hat{x}_{0:k}, u_{0:k}) \in \sigma(y_{0:k-1})$ である。また、

70 5. 状態推定 (17)

$y_{0:k} \perp\!\!\!\perp e_{k+1}$ (定理 5.2.2) および $y_{0:k} \in \sigma(x_0, v_{0:k}, w_{0:k}) \perp\!\!\!\perp w_{k+1}$ (仮定 5.2.1) より (5.25) が成り立つ.

さらに, 評価関数に関して, $x_k = \hat{x}_k + e_k$ および $e_k \sim \mathcal{N}(0, \Sigma_k)$, $\hat{x}_k \perp\!\!\!\perp e_k$ に注意すると,

$$\begin{aligned}\mathbb{E}[\ell(k, x_k, u_k)] &= \mathbb{E}[\ell(k, \hat{x}_k + e_k, u_k)] \\ &= \mathbb{E}[\ell(k, \hat{x}_k, u_k)] + \mathbb{E}[e_k^\top Q_k e_k] \\ &= \mathbb{E}[\ell(k, \hat{x}_k, u_k)] + \text{Trace}(Q_k \Sigma_k) \\ \mathbb{E}[\ell_f(x_{\bar{k}})] &= \mathbb{E}[\ell_f(\hat{x}_{\bar{k}})] + \text{Trace}(\Pi_{\bar{k}} \Sigma_{\bar{k}})\end{aligned}$$

が成り立ち, 両式ともに第二項は u_k に依存しない.

以上をまとめると, 定理 4.1.5 および定理 4.2.4 より, 求める最適制御則は (5.21) で与えられる. 最適コストに関しては, (5.23) の \hat{N}_k が

$$\mathbb{E}[\hat{v}_k] = 0, \mathbb{E}[\hat{v}_k \hat{v}_k^\top] = \hat{N}_k \quad (5.26)$$

を満たすことを確認すればよい. ここで, $w_k \perp\!\!\!\perp e_k \in \sigma(x_0, v_{0:k}, w_{0:k-1})$ より,

$$\mathbb{E}[\hat{v}_k \hat{v}_k^\top] = H_k \tilde{M}_k H_k^\top = \tilde{L}_k \tilde{M}_k^{-1} \tilde{L}_k^\top = \tilde{N}_k - \Sigma_{k+1}$$

であることから求める結果がしたがう. \square

この結果は, 状態推定則と状態フィードバック制御則を個別に設計することを正当化しており, 分離定理 (separation theorem) と呼ばれる.

5.2.3 カルマンフィルタ

前節では, 一段予測問題に対する逐次手法を導出したが, 本節ではフィルタリング問題に対する実用的な手法を紹介する [22]. まず, 線形システムの場合, カルマンフィルタ (Kalman filter) と呼ばれる同様の計算で事後分布が得られる.

定理 5.2.5 (カルマンフィルタ). 定理 5.2.2 の記法および

$$\check{H}_k := \check{L}_k \tilde{M}_k^{-1}, \check{L}_k := \Sigma_k C_k^\top \quad (5.27)$$

$$\check{x}_k := \hat{x}_k + \check{H}_k(y_k - C_k \hat{x}_k) \quad (5.28)$$

$$\check{\Sigma}_k := \Sigma_k - \check{L}_k \tilde{M}_k^{-1} \check{L}_k^\top \quad (5.29)$$

のもと,

$$x_k|_{y_{0:k}=y_{0:k}} \sim \mathcal{N}(\check{x}_k, \check{\Sigma}_k) \quad (5.30)$$

が成り立つ. \triangleleft

証明. (5.19) と同様に

$$\begin{bmatrix} x_k \\ y_k \end{bmatrix} \Bigg|_{y_{0:k-1}=y_{0:k-1}} \sim \mathcal{N} \left(\begin{bmatrix} \hat{x}_k \\ C\hat{x}_k \end{bmatrix}, \begin{bmatrix} \Sigma_k & \check{L}_k \\ \check{L}_k^\top & \tilde{M}_k \end{bmatrix} \right) \quad (5.31)$$

が成り立つので, 定理 2.3.4, 2.3.5 よりただちに導出される. \square

(5.28) はカルマンゲインと呼ばれる \check{H}_k を用いて \hat{x}_k を修正する構造となっている. 以下では, 本節を通じて $L_k = O$ を仮定する. このとき, v_k と w_k は独立であり, $y_k = y_k$ の観測は v_k の分布に影響を与えない. このことは,

$$\hat{x}_{k+1} = A_k \check{x}_k + B_k u_k, \Sigma_{k+1} = A_k \check{\Sigma}_k A_k^\top + N_k \quad (5.32)$$

が成り立つことからも確認できる. この関係を用いると, 定理 5.2.2, 5.2.5 で与えた手続きはアルゴリズム 1 の形に書き直せる. ここで, (5.28) は観測更新式, (5.32) は時間更新式と呼ばれる.

注意 5.2.6. 問題 5.2.3 の LQG 制御問題において $u_k \in \sigma(y_{0:k})$ とすると, 同様に分離定理が成り立ち, 最適制御則は最適状態フィードバックとカルマンフィルタの併合形 $u_k = -K(\Pi_k)\check{x}_k$ で与えられる. \triangleleft

つぎに対象が非線形システム

Algorithm 1 カルマンフィルタ

Require: $\hat{x}_0 := \bar{x}$, $\Sigma_0 := \bar{\Sigma}$, $k = 0$ で初期化

- 1: **for** $k = 0, 1, 2, \dots$ **do**
- 2: 観測値 y_k , u_k を取得
- 3: $\mathbf{h}_k := C_k \hat{x}_k$, $\tilde{M}_k := M_k + C_k \Sigma_k C_k^\top$, $\check{L}_k := \Sigma_k C_k^\top$, $\check{H}_k := \check{L}_k \tilde{M}_k^{-1}$
- 4: $\check{x}_k := \hat{x}_k + \check{H}_k(y_k - \mathbf{h}_k)$, $\check{\Sigma}_k := \Sigma_k - \check{L}_k \tilde{M}_k^{-1} \check{L}_k^\top$
- 5: $\hat{x}_{k+1} := A_k \check{x}_k + B_k u_k$, $\Sigma_{k+1} := A_k \check{\Sigma}_k A_k^\top + N_k$
- 6: $\varphi_{x_k}|_{y_{0:k}=y_{0:k}} \sim \mathcal{N}(\check{x}_k, \check{\Sigma}_k)$, $\varphi_{x_{k+1}}|_{y_{0:k}=y_{0:k}} \sim \mathcal{N}(\hat{x}_{k+1}, \Sigma_{k+1})$
- 7: **end for**

$$\begin{cases} x_{k+1} = f_k(x_k, u_k) + v_k, & k \geq 0, \\ y_k = h_k(x_k) + w_k, \end{cases} \quad (5.33)$$

により与えられる場合を考えよう。ただし $x_k \in \mathbb{R}^{n_x}$, $y_k \in \mathbb{R}^{n_y}$, $u_k \in \sigma(y_{0:k-1})$ とする。この場合は、

- $x_k|_{y_{0:k-1}=y_{0:k-1}}$ が正規分布ではない、
- f_k, h_k が線形関数ではない

ため、定理 5.2.2 のような単純な状態推定はおこなえない。以下では、逐次的に

$$x_k|_{y_{0:k-1}=y_{0:k-1}} \sim \mathcal{N}(\hat{x}_k, \Sigma_k), x_k|_{y_{0:k}=y_{0:k}} \sim \mathcal{N}(\check{x}_k, \check{\Sigma}_k) \quad (5.34)$$

と近似するための代表的手法である**拡張カルマンフィルタ**を紹介する。ここでの方針は、非線形性に対しては線形近似を施すことである。そのために f_k, h_k は微分可能と仮定し、 x に関する非線形関数 $f_k(x, u_k)$, $h_k(x)$ をそれぞれ $x = \check{x}$ まわりで泰ラー展開することで、

$$\begin{aligned} f_k(x, u_k) &\simeq f_k(\check{x}_k, u_k) + \hat{A}_k(x - \check{x}_k), \quad \hat{A}_k := \left. \frac{\partial f_k}{\partial x} \right|_{x=\check{x}_k, u=u_k} \\ h_k(x) &\simeq h_k(\hat{x}_k) + \hat{C}_k(x - \hat{x}_k), \quad \hat{C}_k := \left. \frac{\partial h_k}{\partial x} \right|_{x=\hat{x}_k} \end{aligned}$$

と線形近似する。このとき、 $y_{0:k-1} = y_{0:k-1}$ のもとで $x_k \sim \mathcal{N}(\hat{x}_k, \Sigma_k)$ および $x_k \perp\!\!\!\perp w_k$ ならば、

Algorithm 2 拡張カルマンフィルタ

Require: $\hat{x}_0 := \bar{x}$, $\Sigma_0 := \bar{\Sigma}$, $k = 0$ で初期化

- 1: **for** $k = 0, 1, 2, \dots$ **do**
- 2: 観測値 y_k , u_k を取得
- 3: $\mathbf{h}_k := h_k(\hat{x}_k)$, $\tilde{M}_k := M_k + \hat{C}_k \Sigma_k \hat{C}_k^\top$, $\tilde{L}_k := \Sigma_k \hat{C}_k^\top$, $\hat{C}_k := \frac{\partial h_k}{\partial x} \Big|_{x=\hat{x}_k}$
- 4: $\check{x}_k := \hat{x}_k + \tilde{L}_k \tilde{M}_k^{-1} (y_k - \mathbf{h}_k)$, $\check{\Sigma}_k := \Sigma_k - \tilde{L}_k \tilde{M}_k^{-1} \tilde{L}_k^\top$
- 5: $\hat{x}_{k+1} := f_k(\check{x}_k, u_k)$, $\Sigma_{k+1} := \hat{A}_k \check{\Sigma}_k \hat{A}_k^\top + N_k$, $\hat{A}_k := \frac{\partial f_k}{\partial x} \Big|_{x=\hat{x}_k, u=u_k}$
- 6: $\varphi_{x_k}|_{y_{0:k}=y_{0:k}} \sim \mathcal{N}(\check{x}_k, \check{\Sigma}_k)$, $\varphi_{x_{k+1}}|_{y_{0:k}=y_{0:k}} \sim \mathcal{N}(\hat{x}_k, \Sigma_k)$ により近似
- 7: **end for**

$$\left[\begin{array}{c} x_k \\ h_k(x_k) + w_k \end{array} \right] \Big|_{y_{0:k-1}=y_{0:k-1}} \sim \mathcal{N} \left(\left[\begin{array}{c} \hat{x}_k \\ h_k(\hat{x}_k) \end{array} \right], \left[\begin{array}{cc} \Sigma_k & \Sigma_k \hat{C}_k^\top \\ (\Sigma_k \hat{C}_k^\top)^\top & M_k + \hat{C}_k \Sigma_k \hat{C}_k^\top \end{array} \right] \right) \quad (5.35)$$

であり, $y_{0:k} = y_{0:k}$ のもとで $x_k \sim \mathcal{N}(\check{x}_k, \check{\Sigma}_k)$ かつ $x_k \perp\!\!\!\perp v_k$ ならば,

$$x_{k+1}|_{y_{0:k}=y_{0:k}} \sim \mathcal{N} \left(f_k(\check{x}_k, u_k), \hat{A}_k \check{\Sigma}_k \hat{A}_k^\top + N_k \right) \quad (5.36)$$

である. 以上をまとめると, アルゴリズム 2 が得られる.

注意 5.2.7. 他にも確率分布の時間発展を有限個の粒子の挙動で近似して解析する手法も広く用いられる. 代表例に, 確率変数の平均と分散からシグマポイントと呼ばれる粒子を生成する無香料カルマンフィルタ (Unscented Kalman filter) や, 雑音の正規性は仮定せずベイズ推定と再サンプリングという手法を用いる粒子フィルタ (particle filter) などがある [?]. \diamond

5.3 軌道の MAP 推定と最適制御の双対性

ここでは非線形システムも対象とし, 出力時系列から初期状態や現在の状態のみではなく, 過去の状態時系列を推定する問題を紹介する. 現代制御論において, 可制御性と可観測性の判別条件やグラミアンの定義 ((3.32) と (5.4)) の

形式的な類似性をはじめとして、制御と推定の双対性がいくつか登場したが、本節であつかう最尤スマージング (modal trajectory estimate) と呼ばれる MAP 推定問題は、前章の最適制御問題と明確な双対関係がある。

定理 5.3.1 (最尤スマージングと最適制御の双対性). $x_0, \{v_k\}, \{w_k\}$ はすべて独立であるとし、

$$\begin{cases} x_{k+1} = f_k(x_k) + v_k, \\ y_k = h_k(x_k) + w_k, \end{cases} \quad (5.37)$$

に対して、観測系列 $y_{0:k}$ が $\varphi_{y_{0:k}}(y_{0:k}) > 0$ を満たすとする。このとき事後確率 $\varphi_{x_{0:k}}(x_{0:k}|y_{0:k} = y_{0:k})$ を最大化する状態軌道 $x_{i|k}^{\text{MAP}}$, $i = 0, 1, \dots, k$ は、確定システム

$$x(i+1) = f_k(x(i)) + u(i), \quad i = 0, 1, \dots, k-1 \quad (5.38)$$

に対して評価関数

$$\ell_{\text{ini}}(x(0)) + \sum_{i=0}^{k-1} \ell(i, x(i), u(i)), \quad (5.39)$$

$$\ell_{\text{ini}}(x) := -\log \varphi_{x_0}(x) - \log \varphi_{w_0}(y_0 - h_0(x)) \quad (5.40)$$

$$\ell(i, x, u) := -\log \varphi_{w_{i+1}}(y_{i+1} - h_{i+1}(f_i(x) + u)) - \log \varphi_{v_i}(u) \quad (5.41)$$

を最小化する $x(0), u(0), \dots, u(k-1)$ を求める最適制御問題の最適軌道と一致する。 \triangleleft

証明. この問題に対して、

$$\varphi_{x_{0:k}}(x_{0:k}|y_{0:k} = y_{0:k}) = \frac{\varphi_{y_{0:k}}(y_{0:k}|x_{0:k} = x_{0:k}) \varphi_{x_{0:k}}(x_{0:k})}{\varphi_{y_{0:k}}(y_{0:k})} \quad (5.42)$$

において、独立性の条件から

$$\varphi_{x_{0:k}}(x_{0:k}) = \varphi_{x_0}(x_0) \prod_{i=0}^{k-1} \varphi_{x_{i+1}}(x_{i+1}|x_i = x_i) = \varphi_{x_0}(x_0) \prod_{i=0}^{k-1} \varphi_{v_i}(v_i) \quad (5.43)$$

$$\varphi_{y_{0:k}}(y_{0:k}|x_{0:k} = x_{0:k}) = \prod_{i=0}^k \varphi_{y_i}(y_i|x_i = x_i) = \prod_{i=0}^k \varphi_{w_i}(w_i) \quad (5.44)$$

5.3 軌道の MAP 推定と最適制御の双対性 75

が成り立つ。ただし、 $v_i := x_{i+1} - f_i(x_i)$, $w_i := y_i - h_i(x_i)$ である。したがって、(5.42) で与えられる最大事後確率の対数の最大化を考えると、求める結果がしたがう。 \square

注意 5.3.2. この定理により得られる $x_{0|k}^{\text{MAP}}$ (最大事後確率軌道の初期値) は、 x_0 の MAP 推定値 $\arg \max_{x_0} \varphi_{x_0}(x_0 | y_{0:k} = y_{0:k})$ とは、正規分布の雑音をともなう線形システムの場合を除いて一般に一致しない。 \triangleleft

このように混入した雑音を制御入力とみなし、「大きな駆動雑音 v は実現しにくいことに対応して、制御入力の大きさに応じたコストがかかる」、「大きな観測雑音 w は実現しにくくことに対応して、観測結果 $y_{0:k-1}$ と得られる出力との差に応じたコストがかかる」という対応関係を導入している。具体的には、最尤スマージング問題は、雑音の対数尤度をコストとした制御問題と等価となる。例えば、 $x_0, \{v_k\}, \{w_k\}$ がすべて正規分布にしたがう確率変数であれば、(5.39) は $x_0, h(x_i), u_i$ に関する二次関数となる。

例 5.3.3 (線形システムの最尤スマージング)。 線形システム

$$x_{k+1} = Ax_k + 0.1v_k, x_0 \sim \mathcal{N}(\mu_0, 0.2I) \quad (5.45)$$

$$y_k = Cx_k + w_k \quad (5.46)$$

$$A := \begin{bmatrix} 1.12 & -0.49 \\ 1 & 0 \end{bmatrix}, C := \begin{bmatrix} 1 & 0 \end{bmatrix}, \mu_0 := \begin{bmatrix} 3 \\ 2 \end{bmatrix} \quad (5.47)$$

を考える。ただし、 v_k, w_k は独立同分布で各成分はラプラス分布 $\text{Lap}(0, 1)$ にしたがうとする。このシステムに対して図 5.3a の y が観測値として得られたとする。このとき、二次計画問題

$$\underset{x_0, u}{\text{minimize}} \quad 2.5\|x_0 - \mu_0\|^2 + 10 \sum_{i=0}^{\bar{k}-1} \|u_i\|_1 + \sum_{i=0}^{\bar{k}} \|Cx_i - y_i\|_1 \quad (5.48)$$

$$\text{subj. to} \quad x_{k+1} = Ax_k + u_k \quad (5.49)$$

を解いて得られる図 5.3 に示す軌道が事後確率を最大化する軌道である。 \triangleleft

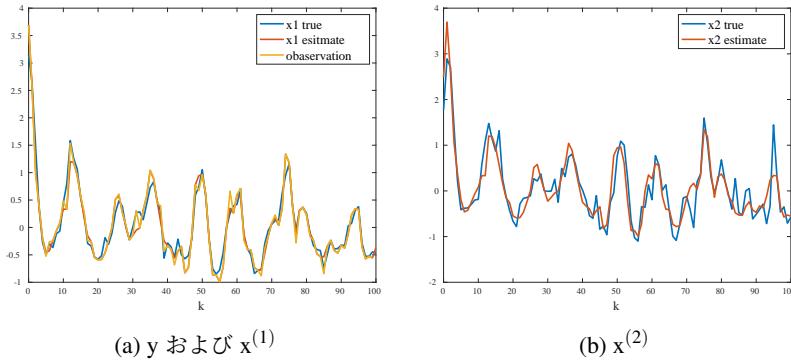


図 5.3: 最尤スムージング

最適制御問題との等価性が示されたため、前章と同様に動的計画法にもとづく構造に注目してみよう。そこで (5.38) に対する到着コスト (arrival cost)

$$V^a(k, x) := \inf_{\substack{x(0), \{u(i)\}_{i=1}^{k-1} \\ \text{s.t. } x(k)=x}} \left\{ \ell_{\text{ini}}(x(0)) + \sum_{i=0}^{k-1} \ell(i, x(i), u(i)) \right\} \quad (5.50)$$

を導入する。明らかに

$$\arg \min_x V^a(k, x) = x_{k|k}^{\text{MAP}} \quad (5.51)$$

が成り立つが、 $i < k$ に対しては $\arg \min_x V^a(i, x) \neq x_{i|k}^{\text{MAP}}$ であることに注意する。また、時間的な入れ子構造より、

$$V^a(0, x) = \ell_{\text{ini}}(x), \quad (5.52)$$

$$V^a(k+1, x) = \min_{x'} \{V^a(k, x') + \ell(k, x, x - f_k(x'))\}, \quad k \in \mathbb{Z}_+ \quad (5.53)$$

が成り立つ。ここで、 $f_k(x') + u = x$ を満たす入力は $u = x - f_k(x')$ であることを用いた。

注意 5.3.4 (Moving horizon 推定). 各時刻 k において、事後確率最大化軌道 $x_{i|k}^{\text{MAP}}, i \in [0, k]$ の最終時刻での値 $x_{k|k}^{\text{MAP}}$ を現時刻での状態推定値として用いる手法を **Moving horizon 推定** と呼ぶ (図 5.1)。このとき、各時刻で定理 5.3.1 の

5.3 軌道の MAP 推定と最適制御の双対性 77

最適制御問題を解いて事後確率最大化軌道を求める必要はない。具体的には、新しく y_k の観測値 $y_k = \hat{y}_k$ を得るたびに、1. (5.53) を用いて $V^a(k, x)$ を求め、2. それを最小化する x が $x_{k|k}^{\text{MAP}}$ である ((5.51) より成り立つ)，という手順を繰り返せばよい。ただし、各時刻において事後確率最大化軌道が求まるわけではないことに注意する。 \triangleleft

さらに、特別な場合として、定理 5.2.2 と同じ部分観測確率システムに対する最尤スムージング問題を考えるとカルマンフィルタとの等価性が導ける。まずは、(5.53) にあらわれる最小化に用いる補題を準備する。

補題 5.3.5. $\begin{bmatrix} x^\top & y^\top \end{bmatrix}^\top$ が正規分布にしたがうとき、 $\max_x \varphi_{(x,y)}(x, y) \propto \varphi_y(y)$ が成り立つ。 \triangleleft

証明. 定理 2.3.5 の記法のもと、 $\varphi_{(x,y)}(x, y) = \varphi_{x|y=y}(x)\varphi_y(y) = \mathcal{N}(x|\hat{x}(y), \hat{\Sigma})\varphi_y(y)$ を最大化する x は $\hat{x}(y)$ である。また、 $\mathcal{N}(\hat{x}(y)|\hat{x}(y), \hat{\Sigma}) = \mathcal{N}(0|0, \hat{\Sigma})$ は y に依存しない定数であることから、求める結果がしたがう。 \square

定理 5.3.6 (線形システムにおけるカルマンフィルタと MH 推定の等価性). 定理 5.2.5 と同じ部分観測確率システムにおいて $L_k = O$ とする。このとき、定理 5.3.1 の事後分布最大化軌道 $x_{i|k}^{\text{MAP}}$ は

$$\check{x}_k (= \arg \max_x \varphi_{x_k|y_{0:k}=y_{0:k}}(x)) = x_{k|k}^{\text{MAP}} \quad (5.54)$$

を満たす。 \triangleleft

証明. (5.51) より、(5.53) で定まる $V^a(k, x)$ が

$$\exp(-V^a(k, x)) \propto \varphi_{x_k|y_{0:k}=y_{0:k}}(x) \quad (5.55)$$

を満たすことを示せばよい。

まず、 $k = 0$ のとき、 $\varphi_{x_0|y_0=y_0}(x) \propto \varphi_{x_0}(x)\varphi_{y_0|x_0=x}(y) = \varphi_{x_0}(x)\varphi_{w_0}(y - h(x))$ および (5.40) より成立する。つぎに、ある k について成立すると仮定すると、

$$\begin{aligned}
& \exp(-V^a(k, x') - \ell(k, x, x - f_k(x'))) \\
&= \exp(-V^a(k, x')) \varphi_{w_{k+1}}(y_{k+1} - h(x)) \varphi_{v_k}(x - f_k(x')) \\
&\propto \varphi_{x_k|y_{0:k}=y_{0:k}}(x') \varphi_{x_{k+1}|x_k=x'}(x) \varphi_{y_{k+1}|x_{k+1}=x}(y_{k+1}) \\
&\propto \varphi_{(x_k, x_{k+1})|y_{0:k+1}=y_{0:k+1}}(x', x)
\end{aligned}$$

が成り立ち、(5.53) より、

$$\exp(-V^a(k+1, x)) \propto \max_{x'} \varphi_{(x_k, x_{k+1})|y_{0:k+1}=y_{0:k+1}}(x', x) \quad (5.56)$$

である。さらに、 $\varphi_{(x_k, x_{k+1})|y_{0:k+1}=y_{0:k+1}}$ は結合正規分布であることから、補題 5.3.5 を適用する (x_k, x_{k+1} が x, y にそれぞれ対応する) と、 k を $k+1$ とした (5.55) が成り立つ。 \square

(5.55) は、線形システムに対しては、カルマンフィルタにより x_k の事後分布を求ることと、Moving horizon 推定により $V^a(k, x)$ を求めることが等価であることを示している。一方で、非ガウス性の雑音を陽に扱う必要がある場合は、例 5.3.3 のような有限時間最適化問題を解くことでよりよい推定値が得られる。

6|| 最適制御 2 (18)

4章では、有限の時間区間上の最適制御問題の解法を概観した。一方で、システム制御の分野では、安定性をはじめとして、十分に時間が経過したのちに対象が望ましい振る舞いをすることを要求することが多い。そこで本章では、確率システムの漸近挙動を概観したのちに、前章までの結果を無限時間区間の場合に拡張する。特に、強化学習や数理ファイナンスにおけるコスト関数に標準的に用いられる減衰係数についても考察する。

6.1 動的計画法

6.1.1 定式化

対象が確率システムの場合、状態変数を特定の点に収束させることができない場合が多く、そもそも平衡点の安定化という制御仕様が有意義ではない。例えば、対象システムが線形の場合であっても、毎時刻雑音 v_k の影響を受け、原点に収束させることができないため、定理 4.2.4において最適制御コストは終端時刻 \bar{k} とともにいくらでも大きくなる。本節では、確定システムの制御とのこうした違いに焦点をあてて、無限区間の最適制御問題の解法を導く。定式化の候補としては、

1. 望ましい定常分布を定め、それに収束させる、
2. パワースペクトル密度の定義 (3.42) と同様に、十分時間が経過した後の時間平均挙動を評価する、

80 6. 最適制御 2 (18)

3. 各時刻のコストを時間的に減衰させる (discounted), などが考えられる。1. については、所望の分布を実現することは理論・実用双方で重要な課題であるものの、定常分布およびそれを達成する制御入力を定めることができがまず困難である。2. の場合の結果は 6.2.3 節で簡単に紹介するが、定常的に望ましい振る舞いをすることを要求する一方で、そこに至るまでの速さやエネルギー効率を考慮することができず、モデル予測制御など近視眼的な制御性能が求められる場合には必ずしも適さない。そこで、本節では過渡応答を適切に考慮しつつ、対象の長時間挙動を適切に評価している 3. の問題設定を採用する。これを採用する利点として、

- ステージコストが漸近的に 0 に収束することが望めない状況においても数学的に意味のある問題設定となる、
- モデルと現実の対象の挙動の差異は時間が経過するにつれて増大すると考えられるため、直近のコストを優先的に評価することは妥当である、
- ダイナミクスとステージコストの時不变性を仮定することで、状態価値関数が状態のみを引数にもつ関数となり、最適制御問題の求解やデータからの学習の効率が改善する、

といったことが挙げられる。

問題 6.1.1 (時不变無限区間最適制御問題). 遷移確率密度関数 Ψ と初期状態 $x_0 \in \text{rv}(\mathbb{X})$ をもつ確率システム、ステージコスト $\ell : \mathbb{X} \times \mathbb{U} \rightarrow \mathbb{R}_+$ 、集合値関数 $U(x) \subset \mathbb{U}$, $x \in \mathbb{X}$ および $\beta \in (0, 1)$ が与えられたとき、

$$u_k \in U(x_k), \forall k \in \mathbb{Z}_+ \quad (6.1)$$

を満たし、期待値

$$J(u) := \mathbb{E} \left[\sum_{k=0}^{\infty} \beta^k \ell(x_k^u, u_k) \right] \quad (6.2)$$

を最小化する因果的な制御入力 u を求めよ. \triangleleft

注意 6.1.2 (極限と期待値 (積分) の交換). 以下の結果の導出において、

$$\mathbb{E} \left[\underbrace{\lim_{n \rightarrow \infty} \sum_{k=0}^n \ell(k, x_k, u_k)}_{\text{確率変数}} \right] = \lim_{n \rightarrow \infty} \mathbb{E} \left[\underbrace{\sum_{k=0}^n \ell(k, x_k, u_k)}_{\text{数列}} \right] \quad (6.3)$$

という条件を多用する。一見、自明に見えるこの極限と期待値演算の交換も、 $\infty - \infty$ 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 の非負値性もその一つであり、他の条件に適宜置き換えることができる。上にも有界であればより強い結果が導けるが、制御工学で標準的な二次形式が除外されるため仮定しない。 \triangleleft

6.1.2 ベルマン方程式

まず、定理 4.1.5、定理 4.1.7 は終端時刻には依存していないため、この無限区間の問題に対しても制御則を $u_k \in \sigma(x_k)$ に限定して議論することができ、ベルマン方程式を導出できる。記法の単純化のため、有限区間の場合と同様に

$$\mathcal{A}_x[V](x, u) := \int_{\mathbb{X}} V(x') \Psi(x'|x, u) dx' = \mathbb{E}[V(x_{k+1}) | x_k = x, u_k = u] \quad (6.4)$$

およびつぎの作用素を定義する。

定義 6.1.3 (ベルマン作用素)。関数 $V : \mathbb{X} \rightarrow \mathbb{R}$ に対して、

$$\mathcal{B}[V](x) := \inf_{u \in U(x)} (\ell + \beta \mathcal{A}_x[V])(x, u) \quad (6.5)$$

により定義される作用素 \mathcal{B} を**ベルマン作用素**と呼ぶ。また、 $i \geq 1$ に対して $\mathcal{B}^{i+1}[V] := \mathcal{B}[\mathcal{B}^i[V]]$ と定義する。 \triangleleft

時不变性を考慮して、改めて (4.26) と同様に状態価値関数

$$V^*(x) := \inf_{u_0 \in \mathcal{U}_a} \mathbb{E}_{x_0=x}^u \left[\sum_{k=0}^{\infty} \beta^k \ell(x_k, u_k) \right], \quad x \in \mathbb{X} \quad (6.6)$$

を定義しよう。ここで、有限区間の場合と異なり、引数に時間変数 k を含まないことに注意する。このとき、補題 4.1.6 と同様に、

$$\inf_{u \in \mathcal{U}_a} J(u) = \mathbb{E}_{x_0}[V^*(x_0)] \quad (6.7)$$

が成り立つ。また、任意の x に対して $V^*(x)$ は有限の値をとるとすると、ベルマン方程式が導出できる。

定理 6.1.4 (ベルマン方程式). 問題 6.1.1 に対して、(6.6) の状態価値関数 V^* はベルマン方程式

$$\mathcal{B}[V] = V \quad (6.8)$$

を満たす。 \triangleleft

証明. 定理 4.1.7 と同様の議論より示される。 \square

定理 4.1.7 と同様に、状態価値関数はベルマン方程式を満たす。しかしながら、有限区間の場合と異なり、終端境界条件が与えられていないため、解は一意ではない。実際、ベルマン方程式は一般に複数の解をもつため、以下の記述において状態価値関数とベルマン方程式の解を混同しないよう注意してほしい。

つぎに、十分性に関する定理 4.2.1 に対応する結果を証明する。定理 4.2.1 の証明では (4.38) が $k = \bar{k}$ において自明に成り立つことを用いて、あとは時刻逆向きに帰納法を適用すればよかった。しかしながら、無限時間区間の場合は、このような自明に成り立つ時刻が存在しないため、別の論法が必要となる。

定理 6.1.5 (最適制御則). 問題 6.1.1において、非負値関数 $\hat{V} : \mathbb{X} \rightarrow \mathbb{R}_+$, $\pi : \mathbb{X} \rightarrow \mathbb{U}$, $\varepsilon \geq 0$ が、

$$\hat{V}(x) + \varepsilon \geq (\ell + \beta \mathcal{A}_x[\hat{V}])(x, \pi(x)), \quad \pi(x) \in U(x) \quad x \in \mathbb{X}, \quad (6.9)$$

を満たすとする。このとき、確定的状態フィードバック $u_k = \pi(x_k)$ は、

$$\mathbb{E}^\pi \left[\sum_{k=0}^{\infty} \beta^k \ell(x_k, u_k) \right] \leq \mathbb{E}_{x_0} [\hat{V}(x_0)] + \frac{\varepsilon}{1 - \beta} \quad (6.10)$$

および

$$\mathbb{E}^\pi \left[\beta^k \hat{V}(x_k) \right] \leq \mathbb{E}_{x_0} [\hat{V}(x_0)] + \frac{\varepsilon}{1-\beta}, \quad k \in \mathbb{Z}_+ \quad (6.11)$$

を満たす。また、 $\hat{V} \geq \mathcal{B}[\hat{V}]$ を満たす任意の非負値関数 \hat{V} および $\varepsilon > 0$ に対して、(6.9) を満たす π が存在し、

$$V^* \leq \hat{V} \quad (6.12)$$

が成り立つ (V^* は (6.6) の状態価値関数)。 \triangleleft

証明。 この π に対して

$$\beta^k \mathbb{E}^\pi [\ell(x_k, \pi(x_k))] \leq \mathbb{E}^\pi \left[\beta^k \hat{V}(x_k) + \beta^k \varepsilon - \beta^{k+1} \hat{V}(x_{k+1}) \right], \quad \forall k \in \mathbb{Z}_+$$

が成り立つ。この不等式の両辺を $k = 0$ から $\bar{k}-1$ まで辺々加えることで、

$$\mathbb{E}^\pi \left[\sum_{k=0}^{\bar{k}-1} \beta^k \ell(x_k, u_k) \right] \leq \mathbb{E} [\hat{V}(x_0)] + \varepsilon \sum_{k=0}^{\bar{k}-1} \beta^k - \beta^{\bar{k}} \mathbb{E}^\pi [\hat{V}(x_{\bar{k}})] \quad (6.13)$$

が成り立ち、 \hat{V} の非負値性から $\bar{k} \rightarrow \infty$ で (6.10) を得る。さらに、 ℓ の非負値性のもと (6.13) の左辺は非負であるため、(6.11) が成り立つ。後半の結果に関して、 $\mathcal{B}[\hat{V}](x) + \varepsilon \geq (\ell + \beta \mathcal{A}_x[\hat{V}])(x, \pi(x))$ を満たす π を選べばよい。さらに、 $x_0 = x$ ($x_0 \sim \delta_x$) および

$$V^\pi(x) := \mathbb{E}_{x_0=x}^\pi \left[\sum_{k=0}^{\infty} \beta^k \ell(x_k, u_k) \right] \quad (6.14)$$

とすると、状態価値関数の定義より $V^*(x) \leq V^\pi(x) \leq \hat{V}(x) + \varepsilon/(1-\beta)$ である。

ここで、 $\varepsilon > 0$ の任意性より (6.12) が成り立つ。 \square

ここで、 $\hat{V} = V^*$ に対して (6.9) を満たす π は、初期状態分布に依存せず $J(\pi) \leq \inf_u J(u) + \varepsilon/(1-\beta)$ を達成するだけでなく、(6.11) および

$$\lim_{k \rightarrow \infty} \beta^k \mathbb{E}^\pi [\ell(x_k, u_k)] = 0 \quad (6.15)$$

の意味での閉ループ安定性を有する。また、(6.12) は、状態価値関数がベルマン方程式の最小の非負値関数解であることを意味する。

無限区間の最適制御問題の下限についても、系 4.2.2 に対応する結果を与える、安定性について考察を加える。

系 6.1.6. 問題 6.1.1 に対して、 \check{V} を

$$\check{V}(x) \leq (\ell + \beta \mathcal{A}_x[\check{V}](x, u)) (x, u), \quad \forall x \in \mathbb{X}, u \in U(x) \quad (6.16)$$

を満たす任意の関数とする。このとき、

$$\lim_{k \rightarrow \infty} \beta^k \mathbb{E}^u [\check{V}(x_k)] = 0 \quad (6.17)$$

を満たす任意の u_* に対して

$$J(u_*) \geq \mathbb{E} [\check{V}(x_0)] \quad (6.18)$$

が成り立つ。 \triangleleft

証明。 (6.13) と同様に、任意の u に対して

$$\mathbb{E}^u \left[\sum_{k=0}^{\bar{k}-1} \beta^k \ell(x_k, u_k) \right] \geq \mathbb{E} [\check{V}(x_0)] - \beta^{\bar{k}} \mathbb{E}^u [\check{V}(x_{\bar{k}})] \quad (6.19)$$

を得る。したがって、(6.17) が成り立つとき、両辺の $\bar{k} \rightarrow \infty$ の極限をとり、極限と積分（期待値）の順序交換（注意 6.1.2）をすることで求める結果が得られる。 \square

後半の解釈を述べるために、任意のベルマン方程式の解 \mathring{V} をとる。このとき、定理 6.1.5 より $\mathring{V} \geq V^*$ である。一方で、 \mathring{V} は (6.16) も満たすため、「 $\mathbb{E} [\mathring{V}(x_0)] > \mathbb{E} [V^*(x_0)]$ となるならば、 $\lim_{k \rightarrow \infty} \beta^k \mathbb{E}^u [\mathring{V}(x_k)] = 0$ は満たさないが $J(u_*) < \mathbb{E} [\mathring{V}(x_0)]$ となる u_* が存在する」ことを (6.18) は意味する。言い換えると、本来の問題にはなかった制約条件 $\lim_{k \rightarrow \infty} \beta^k \mathbb{E}^u [\mathring{V}(x_k)] = 0$ を課すことで増大した性能限界が $\mathbb{E} [\mathring{V}(x_0)] (\geq \mathbb{E} [V^*(x_0)])$ である。

注意 6.1.7 (状態価値関数の下限)。不等式 (6.16) を満たす任意の関数 \check{V} をとる。

このとき、「問題 6.1.1 において評価関数 $J(u_*)$ が有界となる u_* は常に (6.17) を満たす」(たとえば \check{V} が有界) ならば、 $\check{V} \leq V^*$ である ((6.18) において $x_0 \sim \delta_x$ として、 u_* に関する下界を考えればよい)。 \triangleleft

6.1.3 値反復法

状態価値関数 V^* を数値的に求める際、ベルマン方程式 (6.8) を直接的に解くことは困難である。そこで、以下では反復法による解法を紹介する。まず、準備として、ベルマン方程式の単調性と縮小性を確認する。

補題 6.1.8 (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、

$$\mathcal{B}[V_1] \leq \mathcal{B}[V_2] \quad (6.20)$$

が成り立つ。 \triangleleft

証明. 定義 6.1.3 より明らか。 \square

補題 6.1.9 (有界な関数に対するベルマン作用素の縮小性). \mathbb{X} で定義された任意の有界な関数 V_1, V_2 に対して、

$$\|\mathcal{B}^i[V_1] - \mathcal{B}^i[V_2]\|_\infty \leq \beta^i \|V_1 - V_2\|_\infty, \forall i \in \mathbb{Z}_+ \quad (6.21)$$

が成り立つ。 \triangleleft

証明. $c := \|V_1 - V_2\|_\infty$ とすると

$$V_2 - c \leq V_1 \leq V_2 + c \quad (6.22)$$

である。これに (6.20) を適用すると

$$\mathcal{B}[V_2] - \beta c \leq \mathcal{B}[V_1] \leq \mathcal{B}[V_2] + \beta c \quad (6.23)$$

が得られる。これを i 回繰り返すと求める結果がしたがう。 \square

このベルマン作用素の縮小性にもとづいて、つきの反復解法の収束性を示すことができる。

定理 6.1.10 (価値反復法). 問題 6.1.1において, ℓ が有界な関数であるとする.

このとき, (6.6) の状態価値関数 V^* はベルマン方程式の唯一の有界な解であり, 任意の上に有界な関数 V に対して,

$$\lim_{i \rightarrow \infty} \|\mathcal{B}^i[V] - V^*\|_\infty = 0 \quad (6.24)$$

である. \triangleleft

証明. ℓ の有界性より, 状態価値関数 V^* も有界な関数であるため, 補題 6.1.9 より (6.24) がしたがう. また, 一意性に関しては, V を任意の有界な解とするとき, 同様に $\|V - V^*\|_\infty \leq \beta^i \|V - V^*\|_\infty$ であるため, $V = V^*$ が示される. \square

この結果を用いて, ベルマン作用素を繰り返し作用させることにより, ベルマン方程式の解を求める手法を**価値反復法** (value iteration) と呼ぶ. 最後に, この反復中に得られる関数は有限区間最適制御問題の解となる.

補題 6.1.11. 任意の関数 $V : \mathbb{X} \rightarrow \mathbb{R}_+$ および $\bar{k} \geq 1$ に対して,

$$\mathcal{B}^{\bar{k}}[V](x) = \inf_{u_{0:\bar{k}-1} \in \mathcal{U}_a} \mathbb{E}^u \left[\sum_{k=0}^{\bar{k}-1} \beta^k \ell(x_k, u_k) + \beta^{\bar{k}} V(x_{\bar{k}}) \right] \quad (6.25)$$

が成り立つ. 特に, (6.6) の状態価値関数 V^* および任意の $\bar{k} \geq 0$ に対して,

$$\inf_{u \in \mathcal{U}_a} J(u) = \inf_{u_{0:\bar{k}-1} \in \mathcal{U}_a} \mathbb{E}^u \left[\sum_{k=0}^{\bar{k}-1} \beta^k \ell(x_k, u_k) + \beta^{\bar{k}} V^*(x_{\bar{k}}) \right] \quad (6.26)$$

が成り立つ. \triangleleft

証明. $\ell(k, x, u) := \beta^k \ell(x, u)$, $\ell_f(x) := \beta^{\bar{k}} V(x)$, $V'(k, x) := \beta^k \mathcal{B}^{\bar{k}-k}[V](x)$ とおく.

このとき,

$$V'(k, x) = \beta^k \mathcal{B}[\mathcal{B}^{\bar{k}-k-1}[V]] = \inf_{u \in U(x)} (\underbrace{\beta^k \ell(x, u)}_{\ell(k, x, u)} + \underbrace{\beta^{k+1} \mathcal{A}_x [\mathcal{B}^{\bar{k}-k-1}[V]]}_{\int_{\mathbb{X}} V'(k+1, x') \Psi(x'|x, u) dx'})$$

は定理 4.1.7 のベルマン方程式を満たす. また, $V = V^*$ とすると, (6.8) より後半の結果が成り立つ. \square

このように終端コストとして状態価値関数を用いることで、無限区間の問題を有限区間の問題に関連付けることは、モデル予測制御などにおいても用いられる。

6.2 線形制御

6.2.1 安定性

3.2.1節において、線形の確率システムに対する安定性および分散の漸近挙動を調べた。ここでは、時間とともに減衰する係数のもとでの挙動を同様に解析する。

定理 6.2.1 (減衰係数とシューラ安定性). $\beta \in (0, 1)$ とし、線形システム

$$x_{k+1} = Ax_k + v_k, \quad x_0 = x_0 \quad (6.27)$$

において v_k は無相関かつ $\mathbb{E}[v_k] = 0, \text{Var}[v_k] \leq N$ とする。このとき、ある $Q \succ O$ に対して

$$\beta^k \mathbb{E}[x_k^\top Q x_k] \rightarrow 0, \quad \forall x_0 \quad (6.28)$$

ならば $\sqrt{\beta}\rho(A) < 1$ が成り立つ、逆に、 $\sqrt{\beta}\rho(A) < 1$ ならば、任意の $Q \succ O$ に対して (6.28) が成り立つ。 \triangleleft

証明. $\check{x}_k := \sqrt{\beta}^k x_k, \check{A} := \sqrt{\beta} A$ とおくと、

$$\check{x}_{k+1} = \check{A} \check{x}_k + \sqrt{\beta}^{k+1} v_k, \quad \check{x}_0 = x_0 \quad (6.29)$$

が成り立つため、 $\mathbb{E}[\check{x}_k^\top Q \check{x}_k] \geq V(k) := x_0^\top (\check{A}^\top)^k Q \check{A}^k x_0$ である。したがって、(6.28) が成り立つとき、

$$\beta^k \mathbb{E}[x_k^\top Q x_k] = \mathbb{E}[\check{x}_k^\top Q \check{x}_k] \geq \sigma_{\min}(Q) \|\check{A}^k x_0\|^2 \rightarrow 0, \quad \forall x_0 \quad (6.30)$$

から $\rho(\check{A}) < 1$ が成り立つ。

逆に、 $\sqrt{\beta}\rho(A) < 1$ のとき、ある $b \in (\sqrt{\beta}, 1)$ に対して $b\rho(A) < 1$ が成り立つ。このとき、 $\hat{x}_k := b^k x_k$, $\hat{A} := bA$ とおくと、 $\hat{x}_{k+1} = \hat{A}\hat{x}_k + b^{k+1}v_k$ である。また、 $|b| < 1$ に注意し (3.28) を用いると、 $\mathbb{E}[\hat{x}_k^\top Q \hat{x}_k]$ は上に有界であることが示される。最後に、 $\beta^k \mathbb{E}[x_k^\top Q x_k] = (\beta/b^2)^k \mathbb{E}[\hat{x}_k^\top Q \hat{x}_k]$ より (6.28) が成り立つ。

□

安定性とリアプノフ方程式の解の関係に関する定理 3.2.3 に対しても、つきの結果が成り立つ。

定理 6.2.2 (減衰係数とリアプノフ方程式). $\beta \in (0, 1)$ および定理 6.2.1 と同じシステムを考える。このとき、ある $Q \succ O$ に対して

$$\beta A^\top \Pi A + Q = \Pi \quad (6.31)$$

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 $Q \succeq O$ に対してリアプノフ方程式 (6.31) は唯一解 $\bar{\Pi} \succeq Q$ をもち、

$$\mathbb{E} \left[\sum_{k=0}^{\infty} \beta^k x_k^\top Q x_k \right] = x_i^\top \bar{\Pi} x_i + \bar{\gamma} \quad (6.32)$$

を満たす $\bar{\gamma} \geq 0$ が存在する。 ◇

証明. (6.29) に定理 3.2.3, (3.38) を適用すると求める結果がしたがう。ただし、定理 6.2.1 の証明の後半の議論より、 $\beta^k \mathbb{E}[x_k^\top Q x_k] \leq c(\beta/b^2)^k$ を満たす $c > 0$ が存在するため、(6.32) の左辺は有界である。 □

この定理において $v_k = 0$ の確定システムであるとき、 $\beta = 1$ に対しても (6.32) を満たす $\bar{\Pi}$ が存在し $\bar{\gamma} = 0$ となる ((3.38))。一方、 $v_k \neq O$ である確率システムの場合、雑音項の影響が発散しないように $\beta < 1$ が要求され、 $\bar{\gamma} > 0$ である。

6.2.2 無限区間線形二次レギュレータ

本節では、問題 4.2.3 において無限区間としたつきの問題の解を与える。

問題 6.2.3 (無限区間 LQR 問題). 線形確率システム

$$x_{k+1} = Ax_k + Bu_k + v_k, \quad (6.33)$$

$$v_k : \text{i.i.d.}, \mathbb{E}[v_k] = 0, \mathbb{E}[v_k v_k^\top] = N \quad (6.34)$$

を考え、ステージコストが二次形式

$$\ell(x, u) := \begin{bmatrix} x^\top & u^\top \end{bmatrix} \begin{bmatrix} Q & S^\top \\ S & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix}, \quad \begin{bmatrix} Q & S^\top \\ S & R \end{bmatrix} \succ O \quad (6.35)$$

および $U(x) := \mathbb{R}^{n_u}$ とした問題 6.1.1 を無限区間線形二次レギュレータ問題と
いう。 \triangleleft

標準的な現代制御論の講義の主結果である最適制御問題の解 ([1, 定理 8.3])
に対応する結果が、つぎの定理により与えられる。

定理 6.2.4 (無限区間 LQR 問題の最適解). 問題 6.2.3において $\sqrt{\beta}\rho(A - B\bar{K}) < 1$
を満たす \bar{K} が存在すると仮定する。このとき、

$$\tilde{R}(\Pi) := R + \beta B^\top \Pi B, \quad \tilde{S}(\Pi) := S + \beta B^\top \Pi A, \quad \tilde{Q}(\Pi) := Q + \beta A^\top \Pi A \quad (6.36)$$

$$r(\Pi) := \tilde{Q}(\Pi) - \tilde{S}(\Pi)^\top \tilde{R}(\Pi)^{-1} \tilde{S}(\Pi), \quad (6.37)$$

と表記すると、行列リッカチ方程式

$$r(\Pi) = \Pi \quad (6.38)$$

は唯一の正定解 Π^* をもつ。また、

$$V^*(x) := x^\top \Pi^* x + \gamma(\Pi^*), \quad \gamma(\Pi^*) := \frac{\beta}{1-\beta} \text{Trace}(N\Pi^*) \quad (6.39)$$

と定義すると、線形制御則

$$u_k := -K(\Pi^*)x_k, \quad (6.40)$$

$$K(\Pi) := \tilde{R}(\Pi)^{-1} \tilde{S}(\Pi) \quad (6.41)$$

90 6. 最適制御 2 (18)

は最適コスト $\mathbb{E}[V^*(x_0)]$ を達成し,

$$\sqrt{\beta} \rho(A - BK(\Pi^*)) < 1 \quad (6.42)$$

が成り立つ. \triangleleft

減衰係数を導入しているため, 保証される閉ループシステムの安定性もシュー
ア安定よりも弱いことに注意する. 表面的には定理 4.2.4 が自然に拡張されて
いるが, 証明は大きく異なる. まず, 任意の二次形式 $V(x) = x^\top \Pi x + \gamma$ に対し
てベルマン作用素は

$$\mathcal{B}[V](x) = x^\top r(\Pi)x + \beta(\gamma + \text{Trace}(N\Pi)) \quad (6.43)$$

であることが, 定理 4.2.4 の証明と同様の計算

$$\begin{aligned} & \ell(x, u) + \beta \mathbb{E}[V(k, Ax + Bu + v_k)] \\ &= (u + \tilde{R}^{-1}\tilde{S}x)^\top \tilde{R}(u + \tilde{R}^{-1}\tilde{S}x) + x(\tilde{Q} - \tilde{S}^\top \tilde{R}^{-1}\tilde{S})x + \beta(\gamma + \text{Trace}(N\Pi)) \end{aligned} \quad (6.44)$$

により確かめられる. したがって, ステージコストは非有界であるが定理 6.1.10
と同様の結果が成り立つのであれば, 値反復法を適用することにより

$$\Pi_{k+1} := r(\Pi_k), \gamma_{k+1} := \beta(\gamma_k + \text{Trace}(N\Pi_k)) \quad (6.45)$$

で定義される Π_k は定理 6.2.4 のリッカチ方程式の解 Π^* に収束し, 残りの議論
も定理 6.1.4 などからしたがうと期待できる. 実際にこれは正しく, リッカチ
イタレーションと呼ばれる (6.45) はリッカチ方程式に対する反復解法としても
用いられる. これが以下で与える証明の骨子であるが, 細部はやや煩雑かつこ
こでしか用いない論法なので, ひとまず次章に進んでも差し支えない.

それでは, つぎの補題を示したあとに定理 6.2.4 を示す.

補題 6.2.5 (リッカチ方程式の解の存在性). 定理 6.2.4 の仮定のもと, $\Pi_0 := O$, $\gamma_0 := 0$ および (6.45) により $\Pi_k, \gamma_k, k \geq 0$ を定義する. このとき, $\Pi^* := \lim_{k \rightarrow \infty} \Pi_k \succ O$, $\gamma^* := \lim_{k \rightarrow \infty} \gamma_k > 0$ が存在し, リッカチ方程式 $r(\Pi^*) = \Pi^*$ および
 $\gamma(\Pi^*) = \gamma^*$ を満たす. \triangleleft

証明. 補題 6.1.11 と (6.43) から,

$$\inf_{u_{0:i-1}} \mathbb{E}_{x_0=x} \left[\sum_{k=0}^{i-1} \beta^k \ell(x_k, u_k) \right] = x^\top \Pi_i x + \gamma_i \quad (6.46)$$

である. この関係式より $x^\top \Pi_{k+1} x + \gamma_{k+1} \geq x^\top \Pi_k x + \gamma_k, \forall x$, すなわち単調非減少性 $\Pi_k \preceq \Pi_{k+1}$, $\gamma_k \leq \gamma_{k+1}$ がわかる ($x=0$ および $\|x\| \rightarrow \infty$ を考える). また, ステージコストの正定性より $\Pi_1 = Q - S^\top R^{-1} S \succ O$ である. 一方で, 安定化制御則 $\bar{\pi}(x) := -\bar{K}x$ のもとでの安定な線形閉ループシステム $x_{k+1} = (A - B\bar{K})x_k + v_k$ に定理 6.2.2 の後半の結果を適用すると,

$$V^{\bar{\pi}}(x) = \mathbb{E}_{x_0=x} \left[\sum_{k=0}^{\infty} \beta^k \ell(x_k, u_k) \right] = x^\top \bar{\Pi} x + \bar{\gamma} \quad (6.47)$$

を満たす $\bar{\Pi} \succeq O$, $\bar{\gamma} \geq 0$ が存在する. さらに, 明らかに (6.46) は (6.47) により上から抑えられる. ゆえに, 上に有界な単調非減少列 Π_k, γ_k には極限 Π^*, γ^* が存在する. 最後に, 更新則 (6.45) の連続性より, $r(\Pi^*) = \Pi^*$ および $\gamma(\Pi^*) = \gamma^*$ がしたがう. \square

定理 6.2.4 の証明. 非負値関数 $V^*(x) := x^\top \Pi^* x + \gamma^*$ はベルマン方程式の解であり, (6.12) より $V^*(x) \geq V^*(x)$ を満たす. 逆に, (6.46) は $V^*(x)$ 以下なので, $V^*(x) \leq V^*(x)$ が成り立つ. ゆえに V^* は状態価値関数 V^* に一致する. また, (6.44) および定理 6.1.5 から制御則 $\pi^*(x) := -K(\Pi^*)(x)$ の最適性が示される. さらに, この制御則のもとで $x_{k+1} = (A - BK(\Pi^*))x_k + v_k$ は, 任意の x_i に対して

$$\mathbb{E}_{x_0=x_i} \left[\sum_{k=0}^{\infty} \beta^k x_k^\top \underbrace{\begin{bmatrix} I & -K(\Pi^*)^\top \\ S & R \end{bmatrix}}_{\succ O} \underbrace{\begin{bmatrix} Q & S^\top \\ S & R \end{bmatrix} \begin{bmatrix} I \\ -K(\Pi^*) \end{bmatrix}}_{x_k} \right] < \infty \quad (6.48)$$

を満たす. したがって, 定理 6.2.1 より $\sqrt{\beta}\rho(A - BK(\Pi^*)) < 1$ が成り立つ.

最後に一意性を示すため, $\Pi \succ O$ をリッカチ方程式の任意の解とする. ふたび, ベルマン方程式の解 $V(x) := x^\top \Pi x + \gamma(\Pi)$ を構成すると $V \geq V^*$ であり, 補題 6.1.11 から任意の $i \in \mathbb{Z}_+$ に対して

$$\begin{aligned}
V(x) &= \inf_{u_{0:i-1}} \mathbb{E}_{x_0=x}^u \left[\sum_{k=0}^{i-1} \beta^k \ell(x_k, u_k) + \beta^i V(x_i) \right] \\
&\leq \mathbb{E}_{x_0=x}^{\pi^*} \left[\sum_{k=0}^{\infty} \beta^k \ell(x_k, u_k) + \beta^i V(x_i) \right] \\
&= V^*(x) + \beta^i x^\top ((A - BK(\Pi^*))^i)^\top \Pi (A - BK(\Pi^*))^i x
\end{aligned}$$

が成り立つ。ここで制御則の安定性より $i \rightarrow \infty$ の極限で $V(x) \leq V^*(x)$ が成り立ち、求める結果がしたがう。 \square

これで行列リッカチ方程式が解ければ最適フィードバックゲインが求まることが示された。証明の簡単化のため $\Pi_0 = O$ としたが、同様の議論により、任意の半正定行列で初期化してもリッカチイタレーションはリッカチ方程式 (6.38) の一意の正定解に収束することが証明できる。

例 6.2.6. 不安定な線形システム ($\rho(A) = 1.5571$)

$$A := \begin{bmatrix} 0.8 & 0.9 & 0.86 \\ 0.3 & 0.25 & 1 \\ 0.1 & 0.55 & 0.5 \end{bmatrix}, B := \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, v_k, \xi_k \sim \mathcal{N}(0, 1) \quad (6.49)$$

に対する LQR 問題 ($\beta = 0.95$) に対する最適制御則は、リッカチ方程式の解 Π^* を計算することで、 $\pi^*(x) = -K^*x = -[1.0065, 1.6472, 2.2280]x$ と求まる。 $\Pi_0^V = O$, $\Pi_{i+1}^V = r(\Pi_i^V)$ とすると、 $i = 5$ ではじめて安定化制御則 ($\sqrt{\beta}\rho(A - BK(\Pi_5^V)) = 0.9339$) が得られる。また、あらたに $\Pi_0^V = \Pi_{\text{ini}}^V := \Pi_5^V$ と初期化した場合の誤差 $\|\Pi_t^V - \Pi^*\|$ を図 6.1 に示す。 \triangleleft

定理 6.2.4においても、雑音 v_k の大きさは最適制御則に影響しない。したがって、出力フィードバックに関しても同様の分離定理が成り立ち、最適フィルタにより求まる事後平均に (6.41) のゲインを乗じる制御則が最適となる。ここで、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、定数行列に収束する。

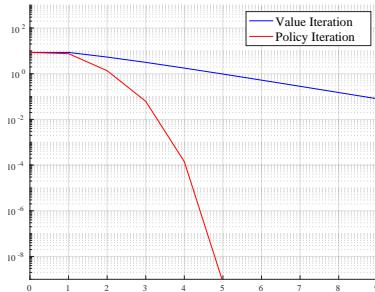


図 6.1: 値反復法と方策反復法

定理 6.2.7 (定常フィルタゲインの安定性). 定理 5.2.2において、各行列は k に依存せず、 $\rho(A - \bar{H}C) < 1$ となる \bar{H} が存在すると仮定する。このとき、

$$\tilde{N}(\Sigma) := N + A\Sigma A^\top, \quad \tilde{M}(\Sigma) := M + C\Sigma C^\top, \quad \tilde{L}(\Sigma) := L + A\Sigma C^\top \quad (6.50)$$

$$r'(\Sigma) := \tilde{N} - \tilde{L}\tilde{M}^{-1}\tilde{L}^\top, \quad (6.51)$$

と定義すると、

$$\Sigma_{k+1} = r'(\Sigma_k), \quad H_k = \tilde{L}(\Sigma_k)\tilde{M}(\Sigma_k)^{-1} \quad (6.52)$$

で与えられる Σ_k はリッカチ方程式

$$\Sigma = r'(\Sigma) \quad (6.53)$$

の一意正定解に収束し、 H_k の収束値 $H := \tilde{L}(\Sigma)\tilde{M}(\Sigma)^{-1}$ に対して $A - HC$ はシュー
ア安定である。 \triangleleft

証明. 定理 6.2.4, 補題 6.2.5 の証明において、 A を A^\top , B を C^\top , $N = O$, $\beta = 1$,
 $\gamma_k = 0$ などと読み替えれば同様の議論が成立する。 \square

6.2.3 【オプション】長時間平均コスト

この節では減衰係数がない $\beta = 1$ の場合の結果を考える。まず確定システムのときは、定理 6.2.4 がそのままの形で成立する。

系 6.2.8 (確定システムに対する無限区間LQR). 問題 6.2.3において, $\beta = 1, N = O$ とする. このとき, $\gamma(\Pi^*) = 0$ とおくと, 定理 6.2.4 はそのまま成立する. \triangleleft

証明. $\beta = 1, N = O$ とおき定理 6.2.4 の証明をたどればよい. \square

なお, $x_0 \sim \mathcal{N}(0, N_i)$ とすると, 最適制御性能は $\text{Trace}(N_i \Pi^*)$ となる. すでに述べたように, この結果を確率システムの場合に拡張すると評価関数が有界とならない. そこで, コストの長時間平均

$$\mathbb{E}^u \left[\limsup_{k \rightarrow \infty} \frac{1}{k} \sum_{i=0}^k \ell(x_i, u_i) \right] \quad (6.54)$$

を考えるとつぎの結果を得る.

系 6.2.9 (長時間平均コストとベルマン不等式). 問題 6.1.1 の確率システムを考え, $\beta = 1$ とする. このとき, 非負値関数 \check{V} および $\varepsilon \geq 0$ が

$$\mathcal{B}[\check{V}] \geq \check{V} + \varepsilon \quad (6.55)$$

を満たすとすると,

$$\lim_{k \rightarrow \infty} \frac{1}{k} \mathbb{E}^u [\check{V}(x_k)] = 0 \quad (6.56)$$

を満たす任意の u に対して (6.54) は ε 以上である. 逆に,

$$\hat{V}(x) + \varepsilon \geq (\ell + \mathcal{A}_x [\hat{V}])(x, \pi(x)), \quad \pi(x) \in U(x) \quad x \in \mathbb{X}, \quad (6.57)$$

を満たす任意の非負値関数 $\hat{V}: \mathbb{X} \rightarrow \mathbb{R}_+$, $\pi: \mathbb{X} \rightarrow \mathbb{U}$, $\varepsilon \geq 0$ に対して, $u(x_k) := \pi(x_k)$ とすると (6.54) は ε 以下である. \triangleleft

証明. 前半の結果は, 仮定よりしたがう不等式 $\mathcal{B}^k[\check{V}] \geq \check{V} + k\varepsilon$ に注意して補題 6.1.11 を適用すると, 任意の u に対して

$$\mathbb{E}_{x_0=x}^u \left[\sum_{i=0}^{k-1} \ell(x_i, u_i) + \check{V}(x_k) \right] \geq \check{V}(x) + k\varepsilon \quad (6.58)$$

が成り立つことから示される. 後半の結果は, $\beta = 1$ のもとで定理 6.1.5 の証明と同様の議論をたどることで示される. \square

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 6.2.10 (長時間平均コストをもつ LQR). 問題 6.2.3において $\rho(A - B\bar{K}) < 1$ を満たす \bar{K} が存在すると仮定する。このとき、(6.36), (6.37) の表記のもと、 $\beta = 1$ とした行列リッカチ方程式 $r(\Pi) = \Pi$ は唯一の正定解 Π^* をもち、 $A - BK(\Pi^*)$ はシューラ安定である。また、(6.54) を最小化する因果的な制御入力 u は

$$u_k := -K(\Pi^*)x_k, K(\Pi) := \tilde{R}(\Pi^*)^{-1}\tilde{S}(\Pi^*) \quad (6.59)$$

で与えられ、最適コストは $\text{Trace}(N\Pi^*)$ である。 \triangleleft

証明. 解の一意存在性およびシューラ安定性は系 6.2.8 すでに確認した。また、直接計算により二次形式 $V(x) := x^\top \Pi^* x$ に対して、 $\mathcal{B}[V](x) = V(x) + \text{Trace}(N\Pi^*)$ が確かめられる。さらに、 $\varepsilon := \text{Trace}(N\Pi^*)$ 、 $\pi(x) := -K(\Pi^*)x$ として系 6.2.9 を用いると、求める結果がしたがう。 \square

最適コストが初期状態に依存しないことからも、長時間平均コストは実質的に定常状態しか評価していないことがわかる。系 6.2.8, 系 6.2.9 ともに出力フィードバックの場合は分離定理が成り立ち、最適フィルタとの併合系が最適となる。この結果は周波数領域の議論からも導出でき、最適 H^2 制御と呼ばれる。

main : 2024/7/13(11:11)

main : 2024/7/13(11:11)

第II部

発展編

main : 2024/7/13(11:11)

7 統計的学習(18)

前章までは、確率分布モデルを所与としたベイズ統計の考え方にもとづいて、確率システムに対する制御系設計論を概観してきた。これに対して、ここからは近年発展の目覚ましい機械学習分野の新技術の活用を試みる。そこで本章では、特に制御系設計に直結することが期待できる関数回帰および特微量抽出、標本生成について、統計的学習の基礎理論と実用的手法を紹介する。

7.1 関 数 回 帰

本節では関数回帰（もしくは教師あり学習（supervised learning）），すなわち未知の関数をその有限個の代入値から推定する問題，およびそれに対する手法を考察する。ここで，推定したい関数に関しての事前情報がなければ明らかにこの問題は不良設定（ill-posed）である。例えば，信号処理におけるシャノンのサンプリング定理では，帯域制限という仮定により対象を滑らかな関数に限定するからこそ，離散的なサンプル値から連続時間信号を復元できる。本節では，こうした推定対象の関数に対する事前情報を明確にすることに留意して，いくつかの代表的な手法を紹介する。

7.1.1 統計推論としての定式化

関数回帰問題を2.3節の枠組みで定式化するために，集合 X （ \mathbb{R}^n などを想定すればよいが，それに限る必要はない）上の実数値関数からなる集合 \mathcal{F} を考え，推定対象の関数は $f \in \text{rv}(\mathcal{F})$ で与えられるとする（図7.1a）。つまり，サ

100 7. 統計的学習 (18)

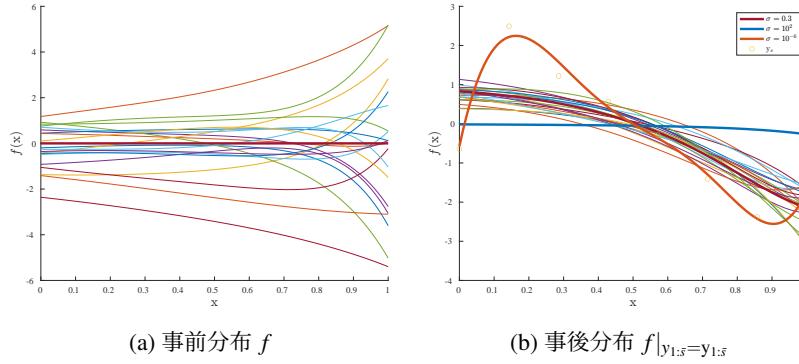


図 7.1: 関数値をとる確率変数の標本関数とモード関数（太線）。

イコロ f を投げると $\varphi_f(f)$ に比例する実現しやすさで $f \in \mathcal{F}$ が一つ定まる（確率密度関数 $\varphi_f(\cdot)$ の具体的な表現や準備方法は後述する）。以下では、（なんらかの x に対する）代入値 $y = f(x)$ の観測結果のもとでの f の事後分布を求めていく（図 7.1b）。

確率過程に図 3.1 のような二つの見方が可能であったのと同様に、確率的に関数 $f \in \mathcal{F}$ が定まる（図 3.1a の標本経路のような視点）定義とは別に、 f は X から $rv(\mathbb{R})$ への関数（代入値 $x \in X$ を固定するごとに $f(x) \in rv(\mathbb{R})$ が定まる図 3.1b のような視点）とみなすこともできる。この準備のもと、既知の評価点 $\{x_s\}_{s=1}^{\bar{s}}$ に対する加法的雑音下の代入値

$$y_{1:\bar{s}} := \begin{bmatrix} y_1 & \cdots & y_{\bar{s}} \end{bmatrix}^\top, y_s := f(x_s) + e_s \quad (7.1)$$

を用いて f の推定をおこなう。ただし、 f, e_1, e_2, \dots は独立であるとする。このとき、 y_s も確率変数であり、

$$\varphi_{y_s}(y_s | f = f) = \varphi_{e_s}(y_s - f(x_s)) \quad (7.2)$$

である。したがって、観測結果 y のもとでの \mathcal{F} 上の事後分布

$$\varphi_f(f | y_{1:\bar{s}} = y_{1:\bar{s}}) \propto \varphi_f(f) \prod_{s=1}^{\bar{s}} \varphi_{e_s}(y_s - f(x_s)) \quad (7.3)$$

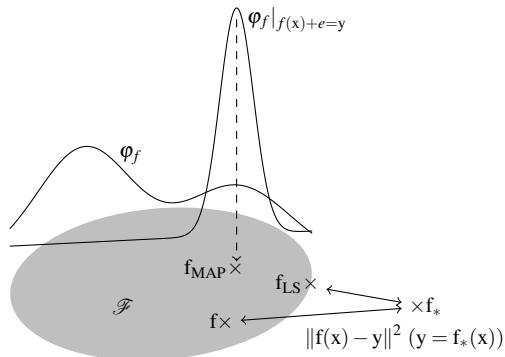


図 7.2: 統計推論および最小二乗法による関数回帰

を分布推定結果とし、(7.3) を最大化する MAP 推定 f_{MAP} などを点推定結果とする（図 7.2）。

ここまで、特別な関数形を仮定せず定式化したが、自由度が高い一方で、事前分布の設定や事後分布の計算が困難となる。そこで、パラメータ $p \in \Theta$ をもつ既知の関数に限定し

$$\mathcal{F} := \{f_p : p \in \Theta\} \quad (7.4)$$

としよう。あとはパラメータを確率変数 $p \in \text{rv}(\Theta)$ とすれば（ p は出現確率・事前知識など好みで解釈すればよい）、 $\varphi_{f_p}(f_p) = \varphi_p(p)$ により $f_p \in \text{rv}(\mathcal{F})$ が設定できる。このように少数のパラメータで表現されるモデルや確率分布を用いる手法をパラメトリック手法、それ以外をノンパラメトリック手法と呼ぶ[†]。同様に、 p, e_1, e_2, \dots は独立、 $y_s := f_p(x_s) + e_s$ とし、事後分布

$$\varphi_p(p | y_{1:\bar{s}} = y_{1:\bar{s}}) \propto \varphi_p(p) \prod_{s=1}^{\bar{s}} \varphi_{e_s}(y_s - f_p(x_s)) \quad (7.5)$$

の計算や最大化、対数尤度

$$\log l_p(p | y_{1:\bar{s}} = y_{1:\bar{s}}) = \sum_{s=1}^{\bar{s}} \log \varphi_{e_s}(y_s - f_p(x_s)) \quad (7.6)$$

[†] 12 章にてノンパラメトリック回帰の代表例であるガウス過程回帰の考え方を紹介する。

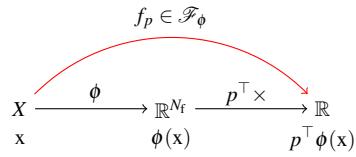


図 7.3: 線形パラメトリック回帰

の最大化などにより推定がおこなえる。

実用上は、図 7.3 のように f_p が p に関して線形、すなわち既知の関数 $\phi: X \rightarrow \mathbb{R}^{N_f}$ を用いて、

$$\mathcal{F}_\phi := \{f_p := \phi^\top p : p \in \mathbb{R}^{N_f}\}, \quad \phi := [\phi_1 \ \dots \ \phi_{N_f}]^\top \quad (7.7)$$

とするモデルは特に重要である。もしも、 $X = \mathbb{R}^n$, $\phi(x) = x$ であれば \mathcal{F}_ϕ は線形関数からなる集合に他ならない（線形回帰）。これでは、候補となる関数が限定的であるため、非線形関数も要素に加えその線形結合まで表現できるようにしたのが (7.7) である。この \mathbb{R}^{N_f} を特徴空間、特徴空間の要素を特徴ベクトル、写像 ϕ を特徴写像と呼ぶ。もしもモデリング対象（関数形）に関する先見情報があるのであれば、それにもとづいて ϕ を構成すればよいが、そうでなければ数学的な扱いやすさを重視することが多い。

例 7.1.1 (多項式回帰). $\phi(x)$ の要素が x の多項式として与えられる関数回帰を多項式回帰と呼ぶ。図 7.1 は、 $X = [0, 1]$ および $\phi(x) = [1, x, \dots, x^9]^\top$, $p \sim \mathcal{N}(0, I)$, $e_s \sim \mathcal{N}(0, \sigma)$, $\sigma = 0.3$ とした場合の標本関数である。 \diamond

演習 7.1. 図 7.1b の太線は $\sigma = 0.3, 10^2, 10^{-6}$ に対する MAP 推定関数である。

この結果の解釈を述べよ。 \diamond

7.1.2 最小二乗法と正則化

7.1.1 節では ϕ_p を準備することで推定対象の関数の \mathcal{F} 上の事前分布を設定し、関数回帰問題を統計推論として定式化した。以下では、確率を用いない標準的な二乗誤差最小化手法と前節の結果の関係を整理する。

(1) 正則化 推定対象の関数が $f_*: X \rightarrow \mathbb{R}$ で与えられ、これを (7.7) の \mathcal{F}_ϕ の要素で近似する問題を考える。そこで、与えられた代入値 $y_s := f_*(x_s)$ に対する重み付き二乗誤差

$$\sum_{s=1}^{\bar{s}} (y_s - f_p(x_s))^2 \quad (7.8)$$

を最小化する $p \in \Theta$ を求めよう。特に、 $f_p := \phi^\top p$ とすると、この評価関数は

$$\phi(x_{1:\bar{s}}) := [\phi(x_1), \dots, \phi(x_{\bar{s}})] \in \mathbb{R}^{N_f \times \bar{s}} \quad (7.9)$$

を用いて二次形式

$$p^\top \phi(x_{1:\bar{s}}) \phi(x_{1:\bar{s}})^\top p - 2p^\top \phi(x_{1:\bar{s}}) y_{1:\bar{s}} + \text{定数} \quad (7.10)$$

とあらわされる。したがって、最小化解

$$p_{LS} := \arg \min_p \sum_{s=1}^{\bar{s}} (y_s - \phi(x_s)^\top p)^2 \quad (7.11)$$

は正規方程式と呼ばれる

$$\phi(x_{1:\bar{s}}) \phi(x_{1:\bar{s}})^\top p = \phi(x_{1:\bar{s}}) y_{1:\bar{s}} \quad (7.12)$$

の解である。ここで、係数行列 $\phi(x_{1:\bar{s}}) \phi(x_{1:\bar{s}})^\top$ が正則でない場合（典型的にはデータ数 \bar{s} が特徴量数 N_f より少ない場合）は劣決定性の問題となり、最小化解は一意に定まらない。同様に、正定でも 0 に非常に近い固有値をもつ場合、最小化解

$$p_{LS} = (\phi(x_{1:\bar{s}}) \phi(x_{1:\bar{s}})^\top)^{-1} \phi(x_{1:\bar{s}}) y_{1:\bar{s}} \quad (7.13)$$

が非常に大きくなったり、データ y の変動に過度に鋭敏となる。逆問題解析における、こうした望ましくない影響を避けるための常套手段として、評価関数に $\|p\|$ のノルムなど適切な項を追加する正則化 (regularization) と呼ばれる手法がある。具体的には、2-ノルムを加える

$$p_{\text{ridge}} := \arg \min_p \sum_{s=1}^{\bar{s}} (y_s - \phi(x_s)^\top p)^2 + \sigma^2 \|p\|^2 \quad (7.14)$$

の最小化をリッジ回帰, 1-ノルムを加える

$$p_{\text{lasso}} := \arg \min_p \sum_{s=1}^{\bar{s}} (y_s - \phi(x_s)^\top p)^2 + \sigma^2 \|p\|_1 \quad (7.15)$$

を **LASSO** (least absolute shrinkage and selection operator), それらを内挿する問題を **エラスティックネット** (Elastic Net) と呼ぶ。いずれも効率的に求解可能な二次計画問題であり、非常に大きなデータ数 \bar{s} の場合にも適用可能な高速アルゴリズムが提案されている [23–25]。こうした正則化手法も、前節の推定の立場からはつぎのように解釈できる。

定理 7.1.2 (事前分布と正則化). $e_s \sim \mathcal{N}(0, \sigma^2)$ のもとで、(7.6) および (7.5) を最大化する $p \in \Theta$ をそれぞれ p_{ML} , p_{MAP} とする。このとき,

- $p_{\text{ML}} = p_{\text{LS}}$
- 事前分布が正規分布 $p \sim \mathcal{N}(0, I)$ のとき, $p_{\text{MAP}} = p_{\text{ridge}}$
- 事前分布がラプラス分布 $\varphi_p(p) = \prod_i \text{Lap}(p_i | 0, 1)$ のとき, $p_{\text{MAP}} = p_{\text{lasso}}$

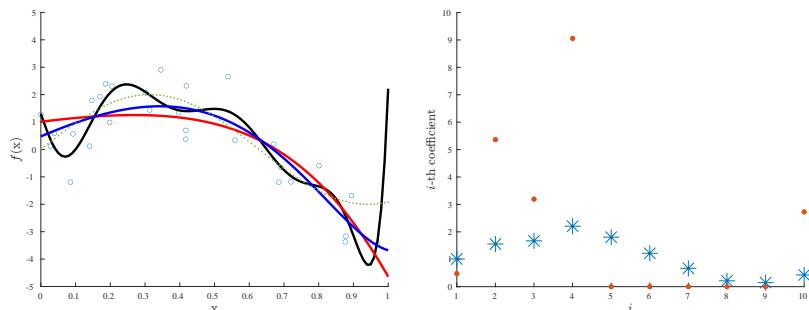
である。 \triangleleft

証明. (7.5) の対数をとればよい。 \square

このように、最小二乗法における正則化項は、統計推論にもとづく関数回帰における事前分布に対応している。また、近似誤差を 2-ノルムで定量化することは観測誤差が正規分布であることを暗に想定している。

例 7.1.3 (LASSO におけるスペース性). $X = [0, 1]$ とし、 $s = 1, \dots, 30$ に対して評価点 x_s は $\text{Uni}(X)$, e_s は $\mathcal{N}(0, 1)$ の独立同分布の乱数により生成し、 $y_s := 2\sin(5x_s) + e_s$ とする (図 7.4a)。また、 $\phi(x) = [1, x, \dots, x^9]^\top$ とし、9 次の多項式回帰をおこなう。まず、正則化がない場合、観測値 y_s に関数を合わせようとしたすぎたため、一般に、この例のように過度に振動する (過学習)。一方、 $\sigma^2 = 0.3$ としたリッジ回帰、LASSO の結果はともに滑らかな関数が得られて

いる。あくまで、「係数ベクトル p は 0 に近い値をとりやすい」という事前知識を取り入れた結果であり、直感的な意味での関数の滑らかさを考慮したわけではないことに注意する。また、得られた p の要素の絶対値を図 7.4b に示す。リッジ回帰においてはすべての要素が非零で比較的小さい値をとっているのに對して、LASSO においては 5 つの要素は 0 であり、非零の要素は比較的大きな大きな値をとっている（正則化がない場合の最大の係数は約 4.5×10^5 ）。これは正規分布とラプラス分布の特性とよく合致している。△



(a) 回帰結果： $2 \sin(5x_s)$ （点線）、最小二乗法（黒）、リッジ回帰（青）、LASSO（赤）
(b) pMAP の要素の絶対値：リッジ回帰（青）、LASSO（赤）

図 7.4: 多項式回帰における事前分布（正則化）の影響。

基底関数の選び方に確立された方法は知られていない。したがって、表現能力を向上するには、できるだけ多くの基底関数をもつことが望ましいが、一般的にデータの次元 n の増加とともに、必要な基底数が増えてしまう問題が生じる。LASSO などに見られる得られる解の要素の多くが 0 となる性質は**スパース性**と呼ばれ、重要な特徴量を客観的に選択し可解釈性を高めることなどに広く用いられる。

注意 7.1.4 (深層学習). ここまで、非線形関数をパラメetrizeるために、既知の基底関数の線形結合を用いた。結果として、パラメータ p の大域的最適解を効率よく求められる。一方で、**深層学習** (deep learning) を始めとする近年のニューラルネットワークを用いた手法では、線形結合ではなく合成関数として関数の集合 \mathcal{F} の表現自由度を非常に高めるだけで、画像処理などの応用例において圧倒的な性能が達成できることが報告されている[†] ◇

(2) カーネルトリック ここでは特徴量数 N_f がデータ数 \bar{s} よりはるかに大きい場合を想定して、特徴写像 ϕ に注目してリッジ回帰問題の構造を具体的に調べてみよう。まず、(7.14) の唯一の最小化解は、

$$p_{\text{ridge}} := (\phi(x_{1:\bar{s}})\phi(x_{1:\bar{s}})^\top + \sigma^2 I)^{-1}\phi(x_{1:\bar{s}})y_{1:\bar{s}} \quad (7.16)$$

と書ける ($y_{1:\bar{s}} \in \mathbb{R}^{\bar{s}}$)。これに逆行列に関する公式を適用すると、つぎのような別表現が得られる。

定理 7.1.5 (カーネルリッジ回帰). カーネル行列と呼ばれる

$$K_\phi(x_{1:\bar{s}}) := \begin{bmatrix} \phi(x_1)^\top \phi(x_1) & \phi(x_1)^\top \phi(x_2) & \cdots & \phi(x_1)^\top \phi(x_{\bar{s}}) \\ \phi(x_2)^\top \phi(x_1) & \phi(x_2)^\top \phi(x_2) & \cdots & \phi(x_2)^\top \phi(x_{\bar{s}}) \\ \vdots & \vdots & \vdots & \vdots \\ \phi(x_{\bar{s}})^\top \phi(x_1) & \phi(x_{\bar{s}})^\top \phi(x_2) & \cdots & \phi(x_{\bar{s}})^\top \phi(x_{\bar{s}}) \end{bmatrix} \quad (7.17)$$

を用いて、(7.14) を最小化する p (すなわち (7.16) の p_{ridge}) は、

$$\phi(x_{1:\bar{s}})(K_\phi(x_{1:\bar{s}}) + \sigma^2 I)^{-1}y_{1:\bar{s}} \in \mathbb{R}^{N_f} \quad (7.18)$$

により与えられる。 ◇

証明. 任意の σ および行列 M に対して $(MM^\top + \sigma^2 I)^{-1}M = M(M^\top M + \sigma^2 I)^{-1}$ が成り立つこと、

[†] 11.2 節では、この背後にある最適化の考え方を紹介する。

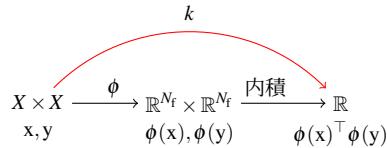


図 7.5: カーネル関数

$$K_\phi(x_{1:\bar{s}}) = \phi(x_{1:\bar{s}})^\top \phi(x_{1:\bar{s}}) \in \mathbb{R}^{\bar{s} \times \bar{s}} \quad (7.19)$$

であることから、求める結果がしたがう。 \square

ここで、 $\phi(x), \phi(y)$ を別々に計算し内積を求める計算量が大きくても、カーネル関数と呼ばれる

$$k(x, y) := \phi(x)^\top \phi(y), x, y \in X \quad (7.20)$$

が容易に計算できる関数であれば、カーネル行列の計算量が大幅に削減でき、これをカーネルトリックと呼ぶ[†]。

例 7.1.6 (多項式カーネル). \mathbb{R}^2 -上の二次関数を表現するために、特徴写像

$$\phi(x) := \begin{bmatrix} 1 & \sqrt{2}(x)_1 & \sqrt{2}(x)_2 & \sqrt{2}(x)_1(x)_2 & (x)_1^2 & (x)_2^2 \end{bmatrix}^\top \in \mathbb{R}^6 \quad (7.21)$$

を導入する ($(x)_i$ は x の第 i 成分)。このとき、 $n = d = 2$ に対して (7.20) の k は

$$k(x, y) = (1 + x^\top y)^d, x, y \in \mathbb{R}^n \quad (7.22)$$

であり、容易に計算できる。より一般的に、 \mathbb{R}^n 上の d 次以下のすべての単項式を適切に係数を定め並べた ϕ に対して、(7.22) が成り立つ。 \triangleleft

SVM

注意 7.1.7 (リプレゼンター定理)。定理 7.1.5 は、最適なパラメータは必ず $p = \phi(x_{1:\bar{s}})p'$, $p' \in \mathbb{R}^{\bar{s}}$ と表現できることを示している。したがって、特徴空間の次

[†] 特徴写像 ϕ ではなくカーネル関数 k に注目した理論体系を 12 章にて紹介する。

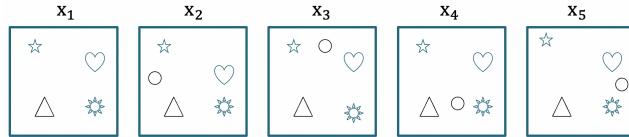


図 7.6: 特徴量

元 N_f がいかに大きくても、データの評価点 $\{x_s\}$ により定まる \bar{s} 次元の特徴写像 $\phi'(\cdot) := \phi(\cdot)^\top \phi(x_{1:\bar{s}})$ により最適な回帰関数は与えられてしまう。また、直接計算により (7.14) は

$$\|y_{1:\bar{s}} - \phi(x_{1:\bar{s}})^\top p\|^2 + \sigma^2 \|p\|^2 = \|y_{1:\bar{s}} - \phi'(x_{1:\bar{s}})^\top p'\|^2 + \sigma^2 \|p'\|_{K_\phi(x_{1:\bar{s}})}^2 \quad (7.23)$$

とかける。これらの関係を定理 12.1.14 と系 12.1.15 で一般化する。 \diamond

7.2 特徴量抽出

与えられたデータ集合に対して、その分布の傾向を特徴づける量を抽出することを、機械学習分野では一般に教師なし学習 (unsupervised learning) と呼ぶ。本節ではその基本的な考え方と手法を概観するが、まずは本書における「特徴量」の考え方をつぎの例を通して説明する。

例 7.2.1. 図 7.6 の画像データにおいて、「星の位置」よりも「丸の位置」を各データの特徴量とすべきと感じるのではないだろうか。その理由としては、「丸の位置は各画像ごとに大きく異なる」や「丸の位置さえわかれば、各画像を精度よく復元できる（たとえば x_1 を背景画像として、適切な位置に丸を加筆する）」などが考えられる。7.2.1 節では前者、7.2.2 節では後者の視点で、それぞれ議論をおこなう。 \diamond

7.2.1 主成分分析

データ集合 $\{x_s\}_{s=1}^{\bar{s}} \subset X$ が与えられたとき、実数値関数 f に対して、標本分散

$$\frac{1}{\bar{s}} \sum_{s=1}^{\bar{s}} \left(f(\mathbf{x}_s) - \frac{1}{\bar{s}} \sum_{s=1}^{\bar{s}} f(\mathbf{x}_s) \right)^2 \quad (7.24)$$

を考える。この値が小さい時、各データ \mathbf{x}_s のかわりに $f(\mathbf{x}_s)$ に注目してもデータ間の差異に関する情報は得られない（例えば、 f が例 7.2.1 の「丸の位置」）。そこで、定数倍の自由度をなくしたうえで、この標本分散を最大化する関数 f を求める。まず、前節同様にパラメータに関して線形な関数 $f_p := \mathbf{p}^\top \phi$ を対象とし、 $\|\mathbf{p}\|^2 \leq 1$ により定数倍の自由度を制限する。 $X = \mathbb{R}^n$, $\phi(\mathbf{x}) = \mathbf{x}$ であれば、ある傾き \mathbf{p} をもつ直線上に射影することにほかならない。

定理 7.2.2 (主成分分析). $\{\mathbf{x}_s\}_{s=1}^{\bar{s}}$ および特徴写像 ϕ が与えられたとき、標本分散

$$\frac{1}{\bar{s}} \sum_{s=1}^{\bar{s}} \left(\mathbf{p}^\top \left(\phi(\mathbf{x}_s) - \frac{1}{\bar{s}} \sum_{s=1}^{\bar{s}} \phi(\mathbf{x}_s) \right) \right)^2 \quad \text{subj. to } \|\mathbf{p}\|^2 \leq 1 \quad (7.25)$$

の最大値 λ および最大化する \mathbf{p} は

$$\frac{1}{\bar{s}} \sum_{s=1}^{\bar{s}} \tilde{\phi}(\mathbf{x}_s) \tilde{\phi}(\mathbf{x}_s)^\top, \quad \tilde{\phi}(\mathbf{x}_s) := \phi(\mathbf{x}_s) - \frac{1}{\bar{s}} \sum_{s=1}^{\bar{s}} \phi(\mathbf{x}_s) \quad (7.26)$$

の最大固有値および対応する固有ベクトルである。 \diamond

証明. 制約条件に対応するラグランジュ乗数を λ とおき、

$$\frac{1}{\bar{s}} \sum_{i=1}^{\bar{s}} \left(\mathbf{p}^\top \tilde{\phi}(\mathbf{x}_s) \right)^2 + \lambda (1 - \mathbf{p}^\top \mathbf{p}) \quad (7.27)$$

とする。これが停留点となるための必要条件として、 \mathbf{p} に関する偏微分を 0 とおくと、ただちに固有方程式

$$\left(\frac{1}{\bar{s}} \sum_{s=1}^{\bar{s}} \tilde{\phi}(\mathbf{x}_s) \tilde{\phi}(\mathbf{x}_s)^\top \right) \mathbf{p} = \lambda \mathbf{p} \quad (7.28)$$

を得る。よって、求める結果がしたがう。 \square

このように標本分散の最大化は、中心化された特徴写像 $\tilde{\phi}$ の標本共分散行列の固有値問題に帰着する。ここでもカーネルトリックの場合と同様に、 $N_f \gg \bar{s}$ の場合を想定して別表現を求めておこう。

補題 7.2.3. 任意の行列 M に対して, $M^\top M$ の 0 でない固有値および対応する固有ベクトルを $\lambda_i, p'_i, i = 1, 2, \dots$, とする. このとき, MM^\top の 0 でないすべての固有値・固有ベクトルは λ_i および $p_i := Mp'_i$ により与えられる. \triangleleft

証明. $MM^\top p = \lambda p$ が成り立つとする. 左から M^\top をかけて $p' := M^\top p \neq 0$ とおくと $M^\top Mp' = \lambda p'$ が成り立ち, p' は $M^\top M$ の固有値 λ に対応する固有ベクトルである. また, M と M^\top を入れ替える, 同じ議論が成り立つ. \square

ここで, $\tilde{\phi}(x_{1:\bar{s}}) := [\tilde{\phi}(x_1), \dots, \tilde{\phi}(x_{\bar{s}})] \in \mathbb{R}^{N_f \times \bar{s}}$ と表記すると, 標本共分散行列 $\tilde{\phi}(x_{1:\bar{s}})\tilde{\phi}(x_{1:\bar{s}})^\top/\bar{s}$ の固有値と固有ベクトルは, $\tilde{\phi}(x_{1:\bar{s}})^\top\tilde{\phi}(x_{1:\bar{s}})/\bar{s}$ の固有値 $\lambda \neq 0$ と対応する固有ベクトル p' を求め, $p := \tilde{\phi}(x_{1:\bar{s}})p'$ とすればよい. さらに, 7.1.1 節 (2) と同様に $K_\phi(x_{1:\bar{s}})$ を K と表記すると,

$$\tilde{\phi}(x_{1:\bar{s}})^\top\tilde{\phi}(x_{1:\bar{s}}) = K - \mathbb{1}^\top K/\bar{s} - K\mathbb{1}/\bar{s} - \mathbb{1}^\top K\mathbb{1}/\bar{s}^2 \quad (7.29)$$

と書き下せる. したがって, $\phi(x_s)$ を具体的に計算することなく, $k(x_s, x_{s'}) := \phi(x_s)^\top\phi(x_{s'})$ を計算すればよい (カーネル主成分分析).

例 7.2.4 (カーネル主成分分析). a \triangleleft

7.2.2 次元削減

ここでは, データの集合が与えられたときに, 適切な精度で復元可能な形で圧縮する問題を考える.

(1) 特異値分解 前節の結果より, 共分散行列

$$\text{Ave}[x] := \frac{1}{\bar{s}} \sum_{s=1}^{\bar{s}} x_s, \text{Var}[x] := \frac{1}{\bar{s}} \sum_{s=1}^{\bar{s}} (x_s - \text{Ave}[x])(x_s - \text{Ave}[x])^\top \quad (7.30)$$

が

$$\text{Var}[x] = \sum_{i=1}^{N_d} \lambda_i v_i v_i^\top, \|v_i\| = 1, \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0 \quad (7.31)$$

と特異値分解できるとき、 $\sum_s \|p^\top(x_s - \text{Ave}[x])\|^2$ は $p = v_1$ により最大化される。これは各 n 次元データ x をスカラー値 $p^\top(x - \text{Ave}[x])$ に情報を圧縮しているとみなすことができる。より一般的に符号化器 (encoder) $E: \mathbb{R}^n \rightarrow \mathbb{R}^m$ により n 次元データを m 次元データに圧縮し、復号器 (decoder) $D: \mathbb{R}^m \rightarrow \mathbb{R}^n$ により復元する、すなわち

$$\frac{1}{\bar{s}} \sum_{s=1}^{\bar{s}} \|D(E(x_s)) - x_s\|^2 \quad (7.32)$$

が小さい値となるような D, E を求めるのことを次元削減と呼ぶ（図 7.7）。上の考察に関連して、 D, E がアファインな場合はつぎの結果が成り立つ [26]。

定理 7.2.5 (主成分分析). 標本共分散行列 $\text{Var}[x]$ が (7.31) と分解できるとき、(7.32) を最小化するアファインな関数の組 (E, D) は、 $V_m := \begin{bmatrix} v_1 & \dots & v_m \end{bmatrix}$ を用いて、

$$E(x) := V_m^\top (x - \text{Ave}[x]) \quad (7.33)$$

$$D(z) := V_m z + \text{Ave}[x] \quad (7.34)$$

で与えられ、最小値は $\sum_{i=m+1}^n \lambda_i$ である。 ◇

つまり m 次元に圧縮するには、第 m 固有ベクトルまで張られる部分空間に射影することが最適であり、 $m = 1$ の場合に定理 7.2.2 が成り立つことからも自然に理解できるであろう（例えば、例 7.2.1において $\text{Ave}[x]$ が背景画像、 E の出力が「丸の位置」）。

非線形関数も考慮する場合、7.1.1 節の関数回帰の場合と同様に、 E_p, D_p という関数形を固定して、(7.32) を p に関して最小化する問題を解くことが考えられる。特にニューラルネットワークによる符号化器・復号器の組は自己符号化器 (autoencoder) と呼ばれ、非線形の次元削減の実用的手法である [27]。

クラスタリング

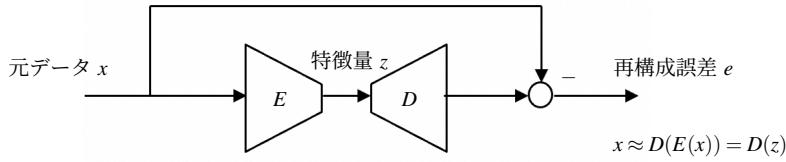


図 7.7: 再構成誤差にもとづく次元削減

7.3 標本の生成と利用

確率変数の統計的性質を調べる際、そのモデルが複雑化するにつれて乱数生成に頼らざるを得ない状況が増える。そこで本節では、標本の生成に関する話題を紹介する。

7.3.1 標本生成

一様分布や正規分布をはじめ、代表的な確率密度関数に対しては効率的な乱数生成法が知られており、様々な計算パッケージで利用できる。一方、任意に与えられた確率密度関数に対して、その乱数生成は必ずしも容易ではない。ここでは、基本的な手法を二つ紹介しておこう。

定理 7.3.1 (逆関数法). 実数上の確率密度関数 $\bar{\varphi}$ の累積分布関数

$$\bar{F}(x) := \int_{-\infty}^x \bar{\varphi}(x') dx' = \mathbb{P}_{\bar{x} \sim \bar{\varphi}}(\bar{x} \leq x) \quad (7.35)$$

を定義する。このとき、 $u \sim \text{Uni}([0, 1])$ に対して、 $x := \bar{F}^{-1}(u)$ の確率密度関数は $\bar{\varphi}$ である。 \triangleleft

証明. $\mathbb{P}(x \leq x) = \mathbb{P}(u \leq \bar{F}(x)) = \bar{F}(x)$ であることからしたがう。 \square

この**逆関数法**により、累積分布関数の逆関数が具体的に計算可能であれば、一様分布にしたがう乱数の生成により所望の分布をもつ乱数が生成できる。一方、つぎの**棄却法**はより一般の確率分布に適用できる。

定理 7.3.2 (棄却法). 実数上の確率密度関数 $\bar{\varphi}$ に対して, $\bar{\varphi}(x) < c\varphi'(x)$, $\forall x$ を満たす任意の定数 $c > 0$ および確率密度関数 φ' をとる. このとき,

1. φ' に従う乱数 y と $\text{Uni}([0, 1])$ の乱数 u を生成する,

2. $\bar{\varphi}(y) \geq uc\varphi'(y)$ ならば $x := y$ として終了し, そうでなければ 1. に戻るとする. このとき x は $\bar{\varphi}$ にしたがう乱数となる. \triangleleft

証明. この手続きにしたがう確率変数 x に対して, 第一巡回で $x \leq x$ が実現する事象は, $y \leq x$ かつ $u \leq \bar{\varphi}(y)/c\varphi'(y)$ であり, その確率は,

$$\mathbb{P}(x \leq x) = \int_{-\infty}^x \varphi'(y) \mathbb{P}\left(u \leq \frac{\bar{\varphi}(y)}{c\varphi'(y)}\right) dy = \frac{\bar{F}(x)}{c}$$

で与えられる (第二巡回に持ち越される確率は $1 - 1/c$ である). この値は第何巡回でも同一であるため, 求める結果がしたがう. \square

ここで, 提案分布と呼ばれる φ' は容易に乱数を生成できる確率密度関数であるとする. この手法は, $\bar{\varphi}$ が $c\varphi'$ により上から抑えられるのであれば, φ' にしたがう乱数生成により $\bar{\varphi}$ にしたがう乱数も生成できる手続きを与えていた. この際, φ' は $\bar{\varphi}$ とはまったく異なってもよいが, c が非常に大きい値となり, 2. で棄却される確率が大きくなり生成効率が悪くなる. 定理 10.2.8において, 確率密度関数の規格化定数が未知の場合に適用可能なメトロポリス・ヘイスティングス法と呼ばれる手法を紹介する.

7.3.2 モンテカルロ法

与えられた確率分布の統計的性質を乱数生成を用いて調べる手法を総じてモンテカルロ法 (Monte Carlo method) と呼ぶ. 以下では, 標本数と得られる情報量の関係に関する代表的な結果を二つ紹介する.

定理 7.3.3. 確率変数 x に対して, $\{y_s\}_{s=1}^{\bar{s}}$ は x のもとで条件付き独立同分布であるとする. このとき, $y_{1:\bar{s}}$ から x を推定する際のフィッシャー情報量行列は

$$\text{FIM}_{y_{1:\bar{s}}|x=x} = \bar{s} \cdot \text{FIM}_{y_1|x=x}, \quad \forall x \quad (7.36)$$

を満たす。 \triangleleft

証明。 仮定より $\varphi_{y_{1:\bar{s}}}(y_{1:\bar{s}}|x=x) = \prod_{s=1}^{\bar{s}} \varphi_{y_1}(y_s|x=x)$ が成り立つことからしたがう。 \square

このようにフィッシャー情報量行列が試行回数に比例して増大する。また、乱数の個数が十分に多い場合の標本平均の性質はつぎの定理により特徴づけられる。

定理 7.3.4 (大数の法則・中心極限定理)。 x_s を平均 μ , 分散 σ^2 をもつ独立同分布にしたがう確率変数とする。このとき, 確率変数 $S_k := \sum_{s=1}^k x_s$ に対して

- S_k/k は μ に確率収束する (**大数の弱法則**)
- S_k/k は μ に概収束する (**大数の強法則**)
- $(S_k - k\mu)/\sigma\sqrt{k}$ は $\mathcal{N}(0, 1)$ に分布収束する (**中心極限定理**)

が成り立つ。 \triangleleft

証明。 弱法則のみ証明を与える。定義より, $\mathbb{E}[\bar{x}_k] = \mu$, $\text{Var}[\bar{x}_k] = \sigma^2/k$ である。

ここで, チェビシェフの不等式 (定理 3.1.13) より $\mathbb{P}(|\bar{x}_k - \mu| \geq \varepsilon) \leq \sigma^2/k\varepsilon^2$ が成り立つ。この不等式で $k \rightarrow \infty$ の極限を考えると確率収束が示される。 \square

具体的に, 確率変数 x および関数 V に対して, $\mathbb{E}[V(x)]$ を計算する状況を考えよう。 $x_s \sim x$ の独立同分布にしたがう確率変数 x_s をとると, $V(x_s)$ も独立同分布にしたがう。したがって, 大数の法則より

$$\frac{1}{\bar{s}} \sum_{s=1}^{\bar{s}} V(x_s) \quad (7.37)$$

は $\mathbb{E}[V(x)]$ に収束する。また, 十分に大きい標本数 \bar{s} に関して (7.37) と期待値 $\mathbb{E}[V(x)]$ の差は正規分布にしたがい, その分散は $1/\bar{s}$ のオーダで小さくなることが中心極限定理の主張である。

注意 7.3.5 (システム制御における仮説検定)。 本書では取り扱わないが, 多数の標本にもとづく仮説検定の理論も, システム制御の分野において安定性解析などに活用されているほか [28], シナリオ最適化という名称で研究されているロバスト性の解析・設計をおこなう手法もある []。 \triangleleft

7.3.3 生成モデル

確率変数 $x \in \text{rv}(X)$ によりモデル化したい現象の独立な標本 $\{x_s\}$ が与えられているとする。このとき、 x の生成モデル (generative model)，すなわち低計算量の標本生成と柔軟な条件付けが可能な数学的モデルの構築を試みる。標本のヒストグラムを作成し、その概形をとらえた関数を $\varphi_x(x)$ とするのが最も直感的だろうか。しかしながら、分布の形状が複雑なとき、その概形をとらえること自体が困難であることに加え、7.3.1 節の手法などで多くの標本を生成することの計算量は大きい。それでは、図 7.7 のような自己符号化器を設計し、デコーダ $D_p : Z \rightarrow X$ を用いて $x = D_p(z)$ としてはどうだろうか。このとき、 D_p への代入は容易であるが、 z の標本生成にはエンコーダを通した標本 $\{E_p(x_s)\}_s$ の分布情報が必要であり、やはり容易ではないことが予想される。

そこで、見方を変えて正規分布のように標本生成が容易な確率変数 z および確率的な生成器 $G_p : Z \rightarrow \text{rv}(X)$ を用いて、 $x = G_p(z) \in \text{rv}(X)$ としてみよう。例えば、

$$G_p(z) = \mu_p(z) + \Sigma_p(z)v, v \sim \mathcal{N}(0, I) \quad (7.38)$$

などにより、 $\varphi_x(x_s | z = z; p)$ を定める。これに対して、

$$\sum_s \log \varphi_x(x_s; p) = \sum_s \log \int_Z \varphi_x(x_s | z = z; p) \varphi(z) dz \quad (7.39)$$

を最大化するという意味で与えられた標本と整合する p を求める。ここで、イエンセンの不等式から、パラメータ q をもち $\int q(z|x; q) dz = 1$ を満たす任意の非負値関数 q に対して、

$$\mathcal{L}_{\text{ELBO}}(p, q; x) := \int q(z|x; q) \log \frac{\varphi_x(x|z=z; p) \varphi_z(z)}{q(z|x; q)} dz \quad (7.40)$$

は $\log \varphi_x(x|z=z; p)$ の上界を与えるが、さらに

$$\log \varphi_x(x; p) = \mathcal{L}_{\text{ELBO}}(p, q; x) + D_{\text{KL}}(q(\cdot|x; q) \| \varphi_z(\cdot|x=x)) \quad (7.41)$$

が成り立つ。したがって、 q に十分な表現能力があれば、 $\mathcal{L}_{\text{ELBO}}(p, q; x)$ は $p^* := \arg \max_p \mathcal{L}_{\text{ELBO}}(p, q; x)$ により最大化される。以上の考察に

116 7. 統計的学習 (18)

より, $\mathcal{L}_{\text{ELBO}}(\mathbf{p}, \mathbf{q}; \mathbf{x})$ の最大化問題をとき, \mathbf{p} を定めることにする. また,

$$\mathcal{L}_{\text{ELBO}}(\mathbf{p}, \mathbf{q}; \mathbf{x}) = \mathbb{E}_{z \sim q(\cdot | \mathbf{x}, \mathbf{q})} [\log \varphi_x(\mathbf{x} | z = z; \mathbf{p})] + D_{\text{KL}}(q(\cdot | \mathbf{x}; \mathbf{q}) \| \varphi_z(\cdot)) \quad (7.42)$$

と書き直せるので, 「標本 \mathbf{x}_s を確率的なエンコーダ q に通した出力結果 z に対して, z をデコーダに通すと \mathbf{x}_s を出す対数尤度が高い (復元誤差が小さい) ことと, z の分布が φ_z から離れすぎないこと」を要求していると解釈できる. この最大化に関する勾配法の計算には, q も (7.38) のような形で与えられるとした変数変換トリック (reparametrization trick) と呼ばれる工夫などが必要であるが, 実装はそれほど難しくない. この手法は, 生成モデルの代表的な手法であり **VAE** (variational autoencoder) と呼ばれる.

また, (7.32) の復元誤差とイエンセン・シャノン情報量

$$D_{\text{JS}}(\mathbf{x} \| \mathbf{y}) := \frac{1}{2} D_{\text{KL}}\left(\varphi_x \middle\| \frac{\varphi_x + \varphi_y}{2}\right) + \frac{1}{2} D_{\text{KL}}\left(\varphi_y \middle\| \frac{\varphi_x + \varphi_y}{2}\right) \quad (7.43)$$

により定量化したエンコーダ出力分布と所望の分布間の距離の和を最小化する確定的なエンコーダ・デコーダを探す問題に, 敵対的学习と呼ばれる手法を活用した手法が敵対的生成ネットワーク (**GAN**; generative adversarial network) である.

これらは2つとも, エンコーダとデコーダの同時最適化が必要であり, 安定な学習が困難な場合も少なくない. これに対して, 雑音を段階的に加えていくという特殊な確率的エンコーダに固定することで, エンコーダ出力分布が自動的に正規分布となるだけでなく, 段階的な雑音除去により安定にデコーダを学習できる拡散モデル (diffusion model) も注目を集めている.

8

システム同定 (19)

6章までは、遷移確率密度関数などの制御対象のモデルが与えられていると仮定し、最適制御、状態推定の理論を概観してきた。本章では、時系列データから動的システムのモデルを構築することを目的とするシステム同定手法の代表的な考え方を紹介する。動的システムは入力信号を引数とし、出力信号を返す（汎）関数とみなすことができることから、システム同定は当然ながら関数回帰問題と密接に関連する[29]。

8.1 最小二乗法

本節では、7.1.2節で検討した最小二乗法を再訪したのち、システム同定への直接的な応用を試みる。

8.1.1 推定器としての最小二乗法

まず、(7.14)のリッジ回帰問題を振り返る。確率変数 $p \sim \mathcal{N}(\bar{p}, \bar{\Sigma})$ および独立同分布の $e_k \sim \mathcal{N}(0, 1)$, $q_k \in \mathbb{R}^n$ に対して,

$$y_k = p^\top q_k + e_k, k = 1, 2, \dots \quad (8.1)$$

が観測できるとする。このとき、 $y_{1:k} = y_{1:k}$ のもとでの p の MAP 推定値は、正則化付き二乗誤差 $\sum_{i=1}^k (y_i - p^\top q_i)^2 + \|p - \bar{p}\|_{\bar{\Sigma}^{-1}}^2$ を最小化する p と一致するのであった。この結果を踏まえて、8.1.1節では、前半で観測値 y_1, y_2, \dots が順番に与えられる場合の推定値の逐次計算法、後半では推定誤差の漸近挙動に関する議論を行なう。

る結果を紹介する。

(1) 逐次最小二乗法 正則化付き二乗誤差の最小化解は、逐次推定の結果と同様につぎの逐次最小二乗法 (Recursive Least Squares Method; RLS) により逐次的に実行することができる。

定理 8.1.1 (逐次最小二乗法). $\bar{p} \in \mathbb{R}^n$, $\bar{\Sigma} \succ O$, $\alpha > 0$ および $q_k \in \mathbb{R}^n$, $y_k \in \mathbb{R}$, $k = 1, 2, \dots$ に対して,

$$\hat{p}_0 := \bar{p}, \Sigma_0 := \bar{\Sigma} \quad (8.2)$$

$$H_k := \frac{\Sigma_{k-1} q_k}{\alpha + q_k^\top \Sigma_{k-1} q_k} \quad (8.3)$$

$$\Sigma_k := \frac{1}{\alpha} (I - H_k q_k^\top) \Sigma_{k-1} \quad (8.4)$$

$$\hat{p}_k := \hat{p}_{k-1} + H_k (y_k - \hat{p}_{k-1}^\top q_k) \quad (8.5)$$

とする。このとき、 \hat{p}_k は

$$\sum_{i=1}^k \alpha^{k-i} (y_i - p^\top q_i)^2 + \alpha^k \|p - \bar{p}\|_{\bar{\Sigma}^{-1}}^2 \quad (8.6)$$

の最小化解である。 \triangleleft

証明. 逆行列補題を用いて直接的に導出できるが、ここではカルマンフィルタとの等価性を示すことで証明する。まず、定理 7.1.2 より、

$$x_k := p, y_k = q_k^\top x_k + w_k, w_k \sim \mathcal{N}(0, \alpha^k), p \sim \mathcal{N}(\bar{p}, \bar{\Sigma}) \quad (8.7)$$

とおくと (8.6) の最小化解は $y_{1:k} = y_{1:k}$ のもとでの x_k の MAP 推定値である。

一方、(8.7) は定理 5.3.6 において状態が時間変化しない線形システム

$$A_k = I, B_k = O, C_k = q_k^\top, N_k = O, L_k = O, M_k = \alpha^k \quad (8.8)$$

とみなすことができる。したがって $\Sigma'_k := \alpha^{-k} \Sigma_k$ とおいた $\Sigma'_0 = \Sigma_0$ および

$$H_k := \frac{\Sigma'_{k-1} q_k}{\alpha^{-k} + q_k^\top \Sigma'_{k-1} q_k}$$

$$\Sigma'_k := (I - H_k q_k^\top) \Sigma'_{k-1}$$

とアルゴリズム 1 を比較すると求める結果がしたがう。 \square

係数が $\alpha < 1$ のとき、小さい i (遠い過去) に対する誤差を小さく評価するため、 α を忘却係数と呼ぶ。分布推定の立場からは、(8.4)において分散行列に相当する Σ_k を $1/\alpha (> 1)$ 倍することは、前ステップまでの観測結果から得られた事後分布の信頼度をあえて下げて事前分布として用いることに相当する。

(2) 一 致 性 定理 8.1.1において、与えられた観測値 $y_{1:k}$ に対して、最小二乗推定値 \hat{p}_k を求める手法を紹介した。具体的には、 $\alpha = 1$ のとき

$$\mathbf{P}_k := \begin{bmatrix} \mathbf{q}_1 & \dots & \mathbf{q}_k \end{bmatrix} \in \mathbb{R}^{n \times k} \quad (8.9)$$

と表記すると、

$$\hat{p}_k = (\mathbf{P}_k \mathbf{P}_k^\top + \bar{\Sigma}^{-1})^{-1} \{ \mathbf{P}_k (y_{1:k} - \mathbb{E}[y_{1:k}]) + \bar{\Sigma}^{-1} \bar{p} \} \quad (8.10)$$

である。ここでは、 p の点推定値として最尤推定値 ($\bar{\Sigma} = +\infty$ の場合の \hat{p}_k) を用いることの妥当性を確認する。最小分散性については、定理 2.3.16においてこの推定量が十分統計量であることを既に示した。一方で、 k が増えるとともに推定誤差が小さくなっていくかについては言及していない。こうした性質を議論するために、つぎの概念を導入する。

定義 8.1.2 (一致推定量)。確率変数 $p, \hat{p}_k \in \text{rv}(\mathbb{R}^n), k = 1, 2, \dots$ を考える。このとき、任意の $p \in \mathbb{R}^n$ に対して、 $p = p$ のもとで \hat{p}_k が p に確率収束するとき、 \hat{p}_k は p の**一致推定量** (consistent estimator) という。 \triangleleft

最小二乗推定の一致性を確認するが、システム同定の文脈で必要となるので、 \mathbf{q}_k が確率変数となる場合に拡張した結果を示しておく。

定理 8.1.3 (最小二乗推定の不偏性・一致性)。 $q_k \in \text{rv}(\mathbb{R}^n)$ および独立同分布の $e_k \in \text{rv}(\mathbb{R})$ に対して、 $y_k = p^\top q_k + e_k, k = 1, 2, \dots$ とし、

$$\mathbb{E}[e_k | q_k = \mathbf{q}_k] = 0, \forall \mathbf{q}_k \quad (8.11)$$

$$\frac{1}{k} \sum_{i=1}^k q_i q_i^\top \xrightarrow{p} \Sigma_q \succ O, \quad (8.12)$$

$$\frac{1}{k} \sum_{i=1}^k q_i e_i \xrightarrow{p} 0 \quad (8.13)$$

が成り立つと仮定する。このとき、 $P_k := \begin{bmatrix} q_1 & \dots & q_k \end{bmatrix} \in \text{rv}(\mathbb{R}^{n \times k})$ の表記のもと、

$$\hat{p}_k := (P_k P_k^\top)^{-1} P_k y_{1:k} \quad (8.14)$$

は p の不偏一致推定量である。 \triangleleft

証明. $y_{1:k} = P_k^\top p + e_{1:k}$ より

$$\hat{p}_k = p + (P_k P_k^\top)^{-1} P_k e_{1:k} \quad (8.15)$$

であることと、(8.11) から $\mathbb{E}[e_{1:k} | P_k = \mathbf{P}_k] = 0, \forall \mathbf{P}_k$ が成り立つことから不偏性がしたがう。つぎに、

$$\hat{p}_k - p = \left(\frac{1}{k} P_k P_k^\top \right)^{-1} \left(\frac{1}{k} P_k e_{1:k} \right) = \left(\frac{1}{k} \sum_{i=1}^k q_i q_i^\top \right)^{-1} \left(\frac{1}{k} \sum_{i=1}^k q_i e_i \right) \quad (8.16)$$

の両辺の確率極限をとることで、(8.12) および (8.13) より一致性が示される。

厳密にはスラツキー (Slutsky) の定理、連続写像定理などを用いるが詳細は割愛する。 \square

積 $q_k^\top p$ の値を介して p の影響を抽出するために、 q_k が 0 に近すぎないことを要求しているのが (8.12) である。また、 y_k から e_k の影響を排除しやすいように、(8.13) は e_k が p_k と類似した成分をあまり含まないことを求めている。例えば、 $q_k \perp\!\!\!\perp e_k$ かつ $\mathbb{E}[e_k] = 0$ であれば (8.11) は成立する。また、 (q_k, e_k) が i.i.d. であれば、大数の弱法則より、

$$\frac{1}{k} \sum_{i=1}^k q_i q_i^\top \xrightarrow{p} \mathbb{E}[q_1 q_1^\top], \quad \frac{1}{k} \sum_{i=1}^k q_i e_i \xrightarrow{p} \mathbb{E}[q_1 e_1] = 0 \quad (\because (8.11)) \quad (8.17)$$

が成り立つ。

8.1.2 システム同定

実数値の入力 u と出力 y をもつ確率システムの同定問題を定式化する。ここで目標は、 u, y の実現値のデータから

$$x_{k+1} = f(x_k, u_k, v_k), \quad y_k = h(x_k) \quad (8.18)$$

が成り立つ状態 x , 雜音 v および f, h を決定することである。そこで、天下り的ではあるが、状態変数は n ステップの出力履歴, m ステップの入力履歴からなる $x_k^\top := [y_k \ \dots \ y_{k-n+1} \ u_{k-1} \ \dots \ u_{k-m}]$ とする。したがって、出力関数は $h(x) = [1 \ 0 \ \dots \ 0] x$ である。さらに、 x_{k+1} の要素のうち x_k と u_k からは定まらない y_{k+1} は x_k, u_k およびそれと独立な v_k の線形結合

$$y_{k+1} = \sum_{i=1}^n a_i y_{k+1-i} + \sum_{i=1}^m b_i u_{k+1-i} + v_k \quad (8.19)$$

で与えられる **ARX モデル** (Auto-Regressive with eXogenous model) を考える。入力をもたない場合 ($b_i = 0$) は AR (AutoRegressive) モデルと呼ばれ、時系列解析で広く用いられる。以下では、パラメータ

$$p := [p_a^\top \ p_b^\top]^\top, \quad p_a := [a_1 \ \dots \ a_n]^\top, \quad p_b := [b_1 \ \dots \ b_m]^\top \quad (8.20)$$

を確率変数とみなし、 $u_{1:k}, y_{1:k}$ の実現値にもとづいて p の推定をする、つまり適切な意味で $\hat{p}_k \rightarrow p$ を満たす $\hat{p}_k \in \sigma(u_{1:k}, y_{1:k})$ を構成する（図 8.2）問題をシステム同定とみなす。

注意 8.1.4. (8.19) は、図 8.1b のように入力 u に雑音 v を加えたものが確定システムに印加される

$$y_{k+1} = \sum_{i=1}^n a_i y_{k+1-i} + \sum_{i=1}^m b_i \bar{u}_{k+1-i}, \quad \bar{u}_k := u_k + v_k \quad (8.21)$$

とは異なり、(8.19) の v はあくまで遷移確率密度関数に対応する確率性である（図 8.1a）。なお、(8.21) のモデルを OE モデルと呼ぶ。 ◇

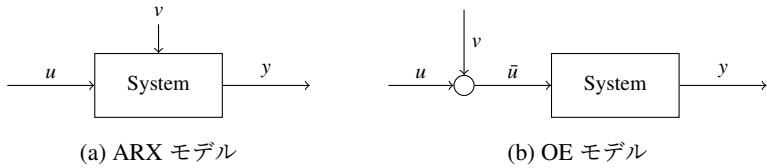


図 8.1: 同定対象のシステム

確率モデル 例:(8.19) システム同定
例：アルゴリズム 3

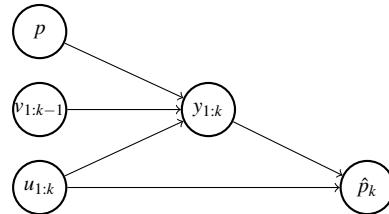


図 8.2: システム同定と確率モデル

さて, (8.19) は

$$q_k^\top := [y_k \ \dots \ y_{k-n+1} \ \ u_k \ \dots \ u_{k-m+1}] \quad (8.22)$$

$$y_k = q_k^\top p + v_k \quad (8.23)$$

とあらわせる. ここで,

$$v_k \sim \mathcal{N}(0, \sigma_v), \text{i.i.d.}, \sigma_v > 0 \quad (8.24)$$

とすると, 前節で考えた推定問題と同一であるため, 結果を直接的に用いることができる. すなわち, 定理 8.1.1 よりアルゴリズム 3 の逐次計算を用いて p の最小二乗推定解 \hat{p}_k が得られ, 適応フィルタの分野などでは **RLS フィルタ** と呼ばれる. なお, $\alpha = 1$ のとき p の事後分布は $\mathcal{N}(\hat{p}_k, \sigma_v \Sigma_k)$ である. 忘却係数 α が 1 に近ければ長い区間での誤差を小さくすることを要求するため, 突発的な外乱の影響に対して頑健である一方, 係数列 p の急激な変化が推定結果になかなか反映されない. 一方, α が小さくなるほど直近の誤差を重点的に評価す

Algorithm 3 忘却係数付き RLS フィルタ

Require: $\hat{x}_0 := \bar{x}$, $\Sigma_0 := \bar{\Sigma}/\sigma_v$, $k = 0$ で初期化

- 1: **repeat**
- 2: 観測値 y_k , u_k を取得し, $q_k^\top := [y_k \dots y_{k-n+1} u_k \dots u_{k-m+1}]$ とする
- 3: $H_k := \frac{\Sigma_{k-1} q_k}{\alpha + q_k^\top \Sigma_{k-1} q_k}$
- 4: $\hat{p}_k := \hat{p}_{k-1} + H_k(y_k - \hat{p}_{k-1}^\top q_k)$
- 5: $\Sigma_k := \frac{1}{\alpha}(I - H_k q_k^\top) \Sigma_{k-1}$
- 6: $k \leftarrow k + 1$
- 7: **until** $k > \bar{k}$

るため推定の速応性がよいが、観測雑音の影響をうけやすい。

例 8.1.5. 確率モデル (8.19) および (8.24) の ARX モデルを採用し, $n = 3$, $m = 2$ および $\bar{p} = 0$, $\bar{\Sigma} = 10^4 \times I$ とする。以下では,

$$(z - 0.5)(z - 0.4)(z - 0.3) = z^3 - a_1 z^2 - a_2 z - a_3, b_1 = 2, b_2 = 1 \quad (8.25)$$

に対して $p^* := [a_1 \ a_2 \ a_3 \ b_1 \ b_2]$, $q_k^\top := [y_k \ y_{k-1} \ y_{k-2} \ u_k \ u_{k-1}]$ のもとで $y_k = q_k^\top p^* + v_k$ を考える。例として, $u_k = 1$, $v_k \sim \mathcal{N}(0, 1)$ および $u_k \sim \mathcal{N}(1, 1)$, $v_k \sim \mathcal{N}(1, 0.01)$ の場合の y の標本経路を図 8.3a に示す。

まず, $u_k \sim \mathcal{N}(0, 1)$, $v_k \sim \mathcal{N}(1, 0.01)$ のもとで a_1 のみ時間的に変化させた場合に、三種類の忘却係数を用いた推定結果を図 8.3b に示す。このとき、 α が小さくなるほど、即応性が向上するものの推定値が振動的となっている。

つぎに、入力に関して $u_k = 1$, $u_k \sim \mathcal{N}(1, 1)$ の二通り、雑音に関して $v_k \sim \mathcal{N}(0, 1)$, $v_k \sim \mathcal{N}(1, 0.01)$ の二通り、の計四通りで標本経路 $u_:$, $y_:$ を生成し、アルゴリズム 3 による \hat{p}_k , Σ_k を計算した。まず、 $u_k = 1$ のとき、 $q_k q_k^\top$ は正則になりえず、(8.12) に対応する条件が成り立たない。結果として、 p_b は真値に漸近しない。一方で、 $u_k \sim \mathcal{N}(1, 1)$ のとき、条件が満たされ $\hat{p} \rightarrow p^*$ および $\hat{\Sigma}_k \rightarrow O$ が成り立つ。また、いずれの場合もシステム雑音 v_k が小さいシステムの方が同定精度は良い。 ◇

7.1 節では、非線形な特徴写像 ϕ を導入することで非線形関数による回帰をおこなう手法を紹介した。(8.22) では右辺に入出力系列を用いたが、ここでも

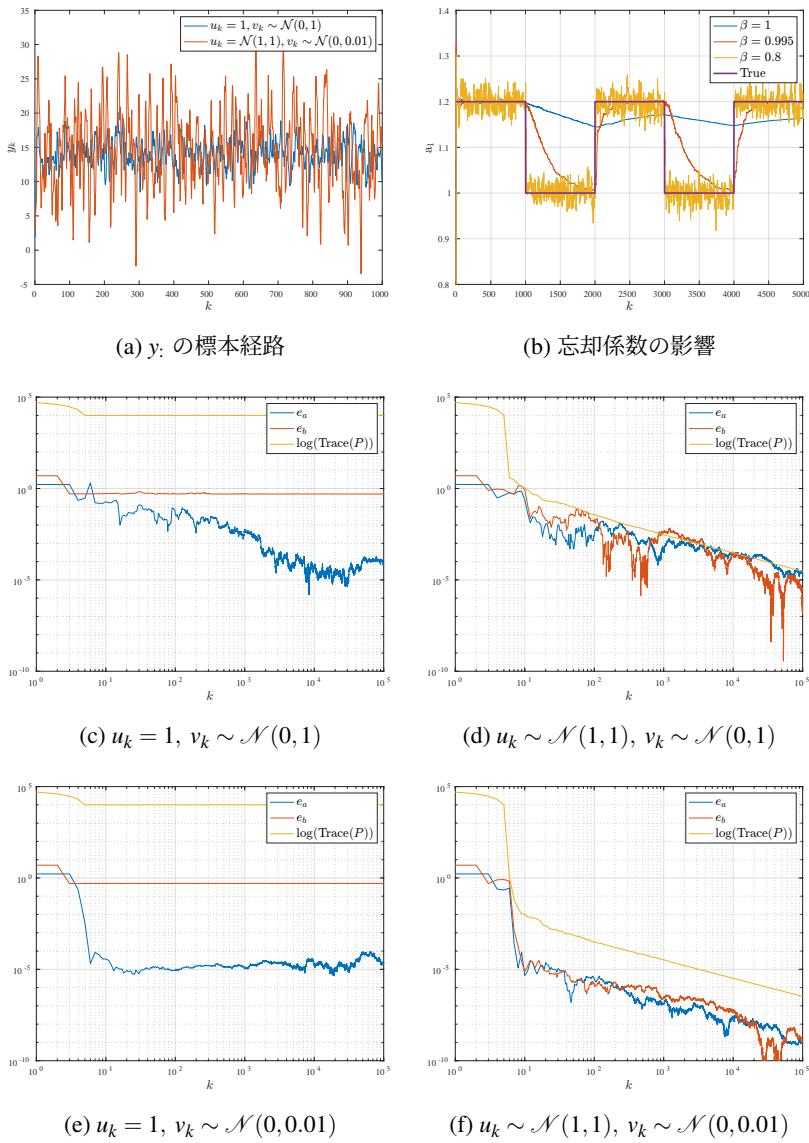


図 8.3: 最小二乗法によるシステム同定結果

非線形関数 ϕ を用いて $q_k := \phi(y_{k-1}, \dots, y_{k-n}, u_{k-1}, \dots, u_{k-m})$ として、同様の議論ができる。この基底関数の選択はボルテラ級数展開とも関係する [30]。

補助変数法

注意 8.1.6. システムをパラメータ p でパラメetrizeし、得られるデータから計算可能な誤差を最小化する手法には様々なものがある。例えば、線形システムの周波数応答が利用可能であるならば、伝達関数を $G_p(z)$ とパラメetrizeし、 $\sum_i \|G_p(e^{-j\omega_i}) - g_i\|$ を最小化すればよい（周波数 ω_i における伝達関数の値 g_i が与えられているとする）。もしくは、 $x_{k+1} = f_p(x_k, u_k)$, $y_k = g_p(x_k)$ で与えられるダイナミクスを仮定し、与えられた $\{(u_k, y_k)\}_k$ を用いて、

$$\mathbb{E}_{x_0} \left[\sum_k \|g_p \circ f_p(\cdot, u_{k-1}) \circ \cdots \circ f_p(\cdot, u_1) \circ f_p(x_0, u_0) - y_k\|^2 \right]$$

を最小化することもできる[†]。ここで、 \circ は関数の合成をあらわす。 ◇

8.2 スペクトル法

本節では、システム同定に関する手法のうち、主として行列の固有値分解にもとづくものを紹介する。具体的には、8.2.1 節では入出力データから状態変数を構成する実現問題に、8.2.2 節では精度をできるだけ保持しつつ状態変数の次元を低くするモデル低次元化問題に焦点をあてる。

8.2.1 正準相関解析と確率実現

(1) 正準相関解析 ここでは、本節で用いる正準相関解析の基本的な考え方と手法を紹介する。二つの確率変数

$$y_+ \in \text{rv}(\mathbb{R}^{\bar{n}}), y_- \in \text{rv}(\mathbb{R}^n)$$

を考え、 y_+ , y_- の平均値を 0, 共分散行列を

[†] こうした合成関数に関する最適化問題の解法を、11.2.2 節で解説する。

$$\mathbb{E} \left[\begin{bmatrix} y_+ \\ y_- \end{bmatrix} \begin{bmatrix} y_+^\top & y_-^\top \end{bmatrix} \right] = \begin{bmatrix} \Sigma_+ & \Sigma_\pm \\ \Sigma_\pm^\top & \Sigma_- \end{bmatrix} \quad (8.26)$$

と表記し, $\Sigma_+ \succ 0, \Sigma_- \succ 0$ を仮定する. ここで, $w_+ := p_+^\top y_+, w_- := p_-^\top y_-$ の相関係数

$$\rho := \frac{\text{cov}(w_+, w_-)}{\sqrt{\text{Var}[w_+]}\sqrt{\text{Var}[w_-]}} = \frac{p_+^\top \Sigma_\pm p_-}{\sqrt{(p_+^\top \Sigma_+ p_+)(p_-^\top \Sigma_- p_-)}} \quad (8.27)$$

を最大にする $p_+ \in \mathbb{R}^{\bar{n}}, p_- \in \mathbb{R}^n$ を見つける問題を考える. これは, y_+, y_- の確率性の共通要素の抽出を試みていると解釈できる. ただし, スカラー倍の自由度をなくすため, 規格化条件

$$p_+^\top \Sigma_+ p_+ = p_-^\top \Sigma_- p_- = 1 \quad (8.28)$$

を課す.

定理 8.2.1. (8.26) の $\Sigma_+, \Sigma_-, \Sigma_\pm$ に対して,

$$\begin{bmatrix} 0 & \Sigma_\pm \\ \Sigma_\pm^\top & 0 \end{bmatrix} \begin{bmatrix} p_+ \\ p_- \end{bmatrix} = \rho \begin{bmatrix} \Sigma_+ & 0 \\ 0 & \Sigma_- \end{bmatrix} \begin{bmatrix} p_+ \\ p_- \end{bmatrix} \quad (8.29)$$

の一般化固有値, すなわち $[p_+^\top \ p_-^\top]^\top \neq 0$ が存在する ρ のうち非負のものを大きい順に $\rho_1 \geq \rho_2 \geq \dots$ のように並べる[†]. このとき, (8.28) のもとで (8.27) の最大値は最大一般化固有値 ρ_1 である. \triangleleft

証明. この制約下での $\rho = p_+^\top \Sigma_\pm p_-$ の最大化問題にラグランジュ未定乗数法を用いるために,

$$\mathcal{L}(p_+, p_-, \lambda_+, \lambda_-) := p_+^\top \Sigma_\pm p_- + \lambda_+(1 - p_+^\top \Sigma_+ p_+) + \lambda_-(1 - p_-^\top \Sigma_- p_-)$$

を考える. 停留条件として, p_-, p_+ に関する偏微分を 0 とおくと

[†] $\bar{n} \leq n$ とすると一般化固有値は $(\rho_1, -\rho_1, \dots, \rho_{\bar{n}}, -\rho_{\bar{n}}, 0, \dots, 0)$ の $\bar{n} + n$ 個である.

$$\Sigma_{\pm} \mathbf{p}_{-} = 2\lambda_{+}\Sigma_{+}\mathbf{p}_{+}, \quad \Sigma_{\pm}^{\top} \mathbf{p}_{+} = 2\lambda_{-}\Sigma_{-}\mathbf{p}_{-} \quad (8.30)$$

であり、左から $\mathbf{p}_+^\top, \mathbf{p}_-^\top$ をそれぞれかけると

$$2\lambda_{+} = 2\lambda_{-} = \mathbf{p}_+^\top \Sigma_{\pm} \mathbf{p}_{-} = \rho \quad (8.31)$$

を得る。したがって、(8.31)を(8.30)に代入することで、最適解 $\mathbf{p}_+, \mathbf{p}_-$ と対応する最大値 ρ は(8.27)を満たすことが示される。□

この ρ_i を第 i 正準相関係数と呼ぶ。具体的な計算法としては、正定値行列 Σ_+, Σ_- を下三角行列とその転置行列の積

$$\Sigma_{+} = LL^{\top}, \quad \Sigma_{-} = MM^{\top} \quad (8.32)$$

に分解（コレスキーディクタント）し、さらに $L^{-1}\Sigma_{\pm}M^{-\top}$ の特異値分解

$$L^{-1}\Sigma_{\pm}M^{-\top} = UDV^{\top}, \quad U^{\top}U = V^{\top}V = I_n, \quad (8.33)$$

$$D = \text{diag}(\rho_i), \quad \rho_1 \geq \rho_2 \geq \dots \geq \rho_n > 0 \quad (8.34)$$

により ρ_i を定めればよく、 $L^{-\top}U, M^{-\top}V$ の各列が一般化固有ベクトル $\mathbf{p}_+, \mathbf{p}_-$ を与える（演習 8.1）。また、正準ベクトルと呼ばれる

$$w_+ := U^{\top}L^{-1}y_+, \quad w_- := V^{\top}M^{-1}y_- \in \text{rv}(\mathbb{R}^n) \quad (8.35)$$

は

$$\mathbb{E}[w_+ w_+^{\top}] = \mathbb{E}[w_- w_-^{\top}] = I_n, \quad (8.36)$$

$$\mathbb{E}[w_+ w_-^{\top}] = D \quad (8.37)$$

を満たす（図 8.4）。このように二つの確率変数間の関係を調べることを正準相関解析（Canonical Correlation Analysis; CCA）と呼び、 y_- が有する y_+ に関する情報は w_- に圧縮でき、 y_+ から得られる y_- に関する情報は w_+ に集約される。また、 D が対角化されていることから、 y_+ の推定のために y_- を $m(< n)$ 次元の情報に圧縮するには、 w_- の第 m 成分までを保持すればよい。

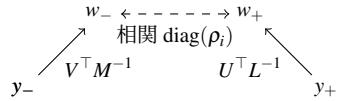


図 8.4: 正準相関解析

演習 8.1. $L^{-\top}U$, $M^{-\top}V$ の第 i 列 p_{i+} , p_{i-} は $p_{i+}^\top \Sigma p_{i+} = p_{i-}^\top \Sigma p_{i-} = 1$ と $p_{i+}^\top \Sigma \pm p_{i-} = \rho_i$ を満たすことを証明せよ. また (8.36), (8.37) を証明せよ. \diamond

(2) 確率実現 \mathbb{R}^{n_y} -値確率過程 $y_:$ の統計的な性質が与えられたときに, マルコフモデルと呼ばれる状態空間モデル

$$x_{k+1} = Ax_k + v_k \quad (8.38)$$

$$y_k = Cx_k + w_k \quad (8.39)$$

を求める問題を考える. ただし, 仮定 5.2.1 と同様に w , v は平均値 0 の白色雑音で, 初期状態 x_0 とは無相関とする.

問題 8.2.2 (確率実現問題). $\sum_{t=-\infty}^{\infty} \|\Lambda(t)\| < \infty$ を満たす $\{\Lambda(t), t = 0, \pm 1, \dots\}$ が与えられるとき, 平均値が $\mathbb{E}[y_k] = 0$, 共分散行列が

$$\mathbb{E}[y_{k+t} y_k^\top] = \Lambda(t), \quad t = 0, \pm 1, \dots$$

を満たす $y_:$ のマルコフモデルを求めよ. \diamond

注意 8.2.3. y の統計的な性質が時刻に依存せず定常状態にあるが, エルゴード性と呼ばれる性質に関しては付録を参照されたい. 確率実現問題では共分散行列 $\{\Lambda(t), t = 0, \pm 1, \dots\}$ が利用可能, 言い換えれば無限長の出力データが利用可能な状況を考えているが, 実際にシステム同定を行う際に利用可能なのは有限の出力標本 $\{y_k, k = 0, 1, \dots, \bar{k}\}$ である. この場合は $\Lambda(\cdot)$ を標本近似することとなる. \diamond

ここで, 状態変数 x のとり方は設計者に委ねられている点が, システム同定の本質的な困難さである. もしも, y の物理的意味が既知であり, 過去の情報

が未来に与える影響がどの値に集約されているかわかるのであれば、その変数を x とすればよいだろう。一方で、 y の統計的性質のみが与えられているときに、過去の情報を集約した変数をどのように構成すればよいだろうか。このように考えると、正準相関解析が状態変数構成の指針を与えることがわかる。

時刻 k を基準に、未来・過去から作られる無限次元ベクトルを

$$y_k^+ := \begin{bmatrix} y_k \\ y_{k+1} \\ \vdots \end{bmatrix}, \quad y_k^- := \begin{bmatrix} y_{k-1} \\ y_{k-2} \\ \vdots \end{bmatrix} \quad (8.40)$$

で定める。さらに問題 8.2.2において y_k^+, y_k^- の相互共分散行列、自己共分散行列

$$\Sigma_{\pm} := \mathbb{E}[y_k^+(y_k^-)^\top] = \begin{bmatrix} \Lambda(1) & \Lambda(2) & \Lambda(3) & \cdots \\ \Lambda(2) & \Lambda(3) & \Lambda(4) & \cdots \\ \Lambda(3) & \Lambda(4) & \Lambda(5) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

$$\Sigma_+ := \mathbb{E}[y_k^+(y_k^+)^\top] = \begin{bmatrix} \Lambda(0) & \Lambda(1)^\top & \Lambda(2)^\top & \cdots \\ \Lambda(1) & \Lambda(0) & \Lambda(1)^\top & \cdots \\ \Lambda(2) & \Lambda(1) & \Lambda(0) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

$$\Sigma_- := \mathbb{E}[y_k^-(y_k^-)^\top] = \begin{bmatrix} \Lambda(0) & \Lambda(1) & \Lambda(2) & \cdots \\ \Lambda(1)^\top & \Lambda(0) & \Lambda(1) & \cdots \\ \Lambda(2)^\top & \Lambda(1)^\top & \Lambda(0) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

は既知である。以降、 y_k^+ と y_k^- に関する正準相関解析を考える。前節と同様に、コレスキー分解 (8.32) で得られる L, M を用いて、特異値分解 (8.33) を行う。ただし、 $n := \text{rank}(\Sigma_{\pm}) < \infty$ を仮定し、 $D = \text{diag}(\rho_1, \dots, \rho_n)$ 、 $1 \geq \rho_1 \geq \dots \geq \rho_n > 0$ 、 $U^\top U = V^\top V = I_n$ である。次に (8.35) に従って、正準ベクトルを

$$w_k^+ = U^\top L^{-1} y_k^+, \quad w_k^- = V^\top M^{-1} y_k^- \quad (8.41)$$

とおき、過去の全情報 y_k^- が未来 y_k^+ に関して有する情報は w_k^- に集約されていると考えられる。便宜上、これをスケーリングした

$$x_k := D^{1/2} w_k^- = D^{1/2} V^\top M^{-1} y_k^- \quad (8.42)$$

を状態変数に採用する。あとは、 y_k の時系列データの標本が得られれば、 x_k の時系列データの標本も得されることになるので、次節で紹介する手法などで $x_{k+1} - Ax_k$ や $y_k - Cx_k$ が適切な意味で小さくなるよう A, C を求めればよい[†]。

8.2.2 モデル低次元化と動的モード分解

(1) 動的モード分解 \mathbb{R}^n -値の状態過程 x_k をもつ線形システム

$$x_{k+1} = Ax_k + Bv_k \quad (8.43)$$

を考える。例えば、偏微分方程式で表現される流体のように空間的に分布した現象を離散化した場合や、各時刻の状態が画像情報として与えられる場合 (n はピクセル数に対応) は、 n が非常に大きな値をとる。しかし、こうした高次元モデルを解析や設計に用いるには、非現実的な計算量が必要となることが多い。そこで本節では、このシステムの性質を保持しつつ、状態変数の次元を削減する**モデル低次元化** (model reduction) を取り組む。もっとも素朴な手法は、状態 x_k の一部を恒等的に 0 とみなすことであろう。具体的には、 $m \ll n$ および $\pi \in \mathbb{R}^{m \times n}$ に対して、 $\Pi := [\pi^\top \underline{\pi}^\top]^\top$ が正則かつ $\pi \underline{\pi}^\top = O$ が成り立つように $\underline{\pi}$ を選ぶと、 $\Pi^{-1} = [\pi^\dagger \underline{\pi}^\dagger]$ であり ($\pi^\dagger = \pi^\top (\pi \pi^\top)^{-1}$)、

$$\begin{bmatrix} \tilde{x}_k \\ x_k \end{bmatrix} := \begin{bmatrix} \pi x_k \\ \underline{\pi} x_k \end{bmatrix}, \quad x_k = \pi^\dagger \tilde{x}_k + \underline{\pi}^\dagger x_k \quad (8.44)$$

$$\begin{bmatrix} \tilde{x}_{k+1} \\ x_{k+1} \end{bmatrix} = \begin{bmatrix} \pi A \pi^\dagger & \pi A \underline{\pi}^\dagger \\ \underline{\pi} A \pi^\dagger & \underline{\pi} A \underline{\pi}^\dagger \end{bmatrix} \begin{bmatrix} \tilde{x}_k \\ x_k \end{bmatrix} + \begin{bmatrix} \pi B \\ \underline{\pi} B \end{bmatrix} v_k \quad (8.45)$$

[†] 付録??において、 A, C などのより具体的な表現を与えるとともに、カルマンフィルタとの等価性を示す。

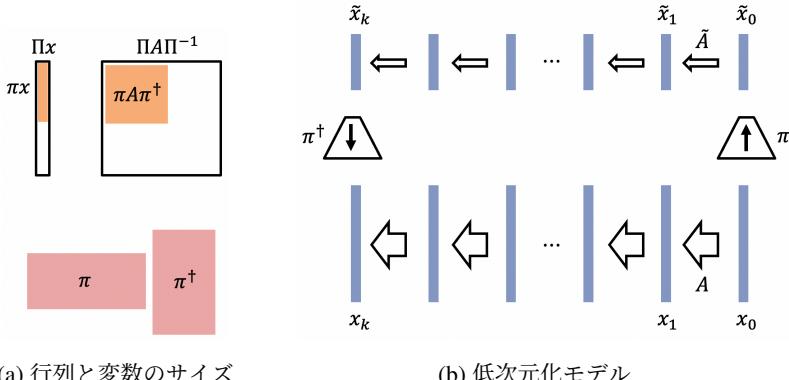


図 8.5: 射影によるモデル低次元化

と書ける。このとき、近似的に

$$x_k \in \text{Im } \pi^\top \quad (8.46)$$

が成り立つならば、 $x_k = 0$ であり、低次元化された状態変数 \tilde{x}_k の時間発展

$$\tilde{x}_{k+1} = (\pi A \pi^\dagger) \tilde{x}_k + \pi B v_k \quad (8.47)$$

のみ追跡すれば、 $x_k = \pi^\dagger \tilde{x}_k$ により x_k が復元できる（図 8.5）。こうした考え方で低次元モデルを得る手法を一般にペトロフ・ガレルキン（Petrov-Galerkin）射影と呼ぶ。例えば、可制御性行列 $C_n := [B \ BA \ \dots \ BA^{n-1}]$ に注目するクリロフ（Krylov）法など、モデル低次元化理論では、与えられた A に対して、条件を満たす π を構成する様々な手法が得られている [31]。

以下では $B = 0$ （外生入力なし）とし、行列 A ではなく実現値の時系列 $\{x_k\}_{k=0}^{\bar{k}}$ が与えられたときに、低次元モデルを求める問題を考えよう。最尤推定や最小二乗法の観点から $\sum \|x_{k+1} - Ax_k\|^2$ の最小化解 A_{LS} の性質を調べる。具体的には、

$$X := \begin{bmatrix} x_0 & x_1 & \dots & x_{\bar{k}-1} \end{bmatrix}, Y := \begin{bmatrix} x_1 & x_2 & \dots & x_{\bar{k}} \end{bmatrix} \quad (8.48)$$

および X の特異値分解

$$\mathbf{X} = \mathbf{U}\mathbf{D}\mathbf{V}^\top, \mathbf{U}^\top\mathbf{U} = \mathbf{V}^\top\mathbf{V} = \mathbf{I} \quad (8.49)$$

を用いて、最小化解は $A_{LS} := \mathbf{Y}\mathbf{X}^\dagger$, $\mathbf{X}^\dagger = \mathbf{V}\mathbf{D}^{-1}\mathbf{U}^\top$ と書ける。

ところで、実応用例においては、データのほとんどが \mathbb{R}^n よりもはるかに小さい集合上に分布していることが多い。したがって、 \mathbf{X} の特異値分解を用いて $x_k \in \text{Im } \mathbf{U}$ とする低次元化手法[†]が有効であり、**固有直交分解** (Proper Orthogonal Decomposition; POD) と呼ばれる。ここで、 \mathbf{U} は大規模密行列 $\mathbf{X}\mathbf{X}^\top$ の非零固有値に対する固有ベクトルを並べた行列であるが、補題 7.2.3 などにより比較的効率よく計算できる。また、 \mathbf{X} の主成分分析だけでなく、データ $\{\mathbf{x}_k\}_k$ の順序も活用することで、つぎのように更なる低次元化がおこなえる。

定理 8.2.4 (動的モード). $\tilde{\mathbf{A}} := \mathbf{U}^\top \mathbf{Y} \mathbf{V} \mathbf{D}^{-1}$ の 0 でない固有値および対応する固有ベクトルを $\lambda_i, \zeta_i, i = 1, 2, \dots, m$ とする。このとき、 $A_{LS} := \mathbf{Y}\mathbf{X}^\dagger$ の 0 でないすべての固有値・固有ベクトルは $\lambda_i, \eta_i := \mathbf{Y} \mathbf{V} \mathbf{D}^{-1} \zeta_i$ により与えられる。 \triangleleft

証明. $\mathbf{X}^\dagger = \mathbf{V}\mathbf{D}^{-1}\mathbf{U}^\top$ に注意すると、

$$A_{LS}(\mathbf{Y} \mathbf{V} \mathbf{D}^{-1} \zeta_i) = (\mathbf{Y} \mathbf{V} \mathbf{D}^{-1}) \underbrace{\mathbf{U}^\top (\mathbf{Y} \mathbf{V} \mathbf{D}^{-1})}_{=\tilde{\mathbf{A}}} \zeta_i = \lambda_i (\mathbf{Y} \mathbf{V} \mathbf{D}^{-1} \zeta_i) \quad (8.50)$$

である。逆に、 $\lambda v = A_{LS}v = \mathbf{Y} \mathbf{V} \mathbf{D}^{-1} \mathbf{U}^\top v$ とすると、

$$\lambda \mathbf{U}^\top v = \mathbf{U}^\top \mathbf{Y} \mathbf{V} \mathbf{D}^{-1} (\mathbf{U}^\top v) = \tilde{\mathbf{A}}(\mathbf{U}^\top v) \quad (8.51)$$

から、 $\mathbf{U}^\top v \neq 0$ は $\tilde{\mathbf{A}}$ の固有ベクトルである。 \square

なお、 $\mathbf{U} \zeta_j$ は $\mathbf{U} \tilde{\mathbf{A}} \mathbf{U}^\top$ (図 8.5a のように (1,1) ブロック以外を O とした行列をもとの座標にもどした行列) の固有ベクトルであり、 A_{LS} の固有ベクトルを $\text{Im } \mathbf{U}$ に射影したものに一致する ($\mathbf{U} \zeta_j = \lambda_j \mathbf{U} \tilde{\mathbf{A}} \zeta_j = \lambda_j \mathbf{U} \mathbf{U}^\top \eta_j$)。定理 8.2.4 から、

[†] 雑音により駆動されるデータに関しては、この低次元化は状態空間を可到達性の高い部分空間に限定してシステム同定をおこなうことに等しく、この考え方は非線形システムにも拡張できることを??節にて示す。

$$x_k \approx A_{LS}^k x_0 = \sum_{i=1}^m (a)_i \lambda_i^k \eta_i, \quad a := [\eta_1 \ \dots \ \eta_m]^\dagger x_0 \quad (8.52)$$

と近似する一連の考え方は動的モード分解 (Dynamic Mode Decomposition; DMD) と呼ばれる。

(2) クーパマン作用素 前節の内容は線形のダイナミクスに特化した議論であるが、7.1 節と同様に非線形拡張をおこなうことができる。既知の関数 $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^{N_f}$ を用いて、

$$\bar{x}_k := \phi(x_k) := [\phi_1(x_k) \ \dots \ \phi_{N_f}(x_k)]^\top \quad (8.53)$$

が線形システムにしたがう、つまり

$$\phi(x_{k+1}) = A_f \phi(x_k) \quad (8.54)$$

と仮定すれば同様の手順で、 A_f とその低次元表現を求めることができる。本節では、この計算がある作用素の近似ともみなせることを説明する。

例 8.2.5 (線形システムの観測作用素の時間発展)。初期状態 x_0 と時刻 k での観測値 y_k の関係が $y_k = CA^k x_0$ と書けるとする。通常は、時刻 k とともに状態変数が線形システム $x_{k+1} = Ax_k$ により時間発展し、各時刻における観測操作は行列 C の乗算に対応すると捉えるだろう。一方で、状態変数は x_0 から変化せず、時刻 k における観測操作が左から $C_k := CA^k$ をかけることであるとみなすこともできる。言い換えると、観測操作は右から A をかけるという作用 $\mathcal{K}[C] := CA$ にしたがって時間発展し、 $C_{k+1} = \mathcal{K}[C_k]$ および $y_k = \mathcal{K}^k[C]x_0$ である。 ◇

この観測操作が時間発展するという見方は、図 8.6 のようにあらわされ、つぎのように非線形システムの場合に拡張できる。

定義 8.2.6 (クーパマン作用素)。 f を \mathbb{X} 上の関数とする。このとき、 $L^\infty(\mathbb{X})$ を定義域とする関数 \mathcal{K} が

$$(\mathcal{K}[h])(\cdot) = h(f(\cdot)), \quad h \in L^\infty(\mathbb{X}) \quad (8.55)$$

を満たすとき、 \mathcal{K} を f のクーパマン作用素 (Koopman operator) と呼ぶ。 ◇

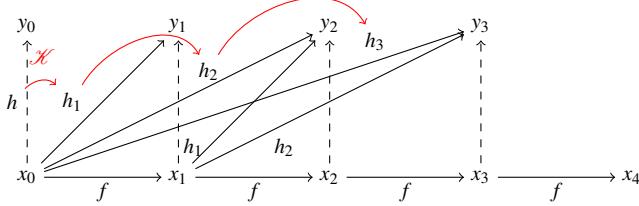


図 8.6: クーブマン作用素

例 8.2.5において、 $(\mathcal{K}[h])(x) := h(Ax)$ は $f(x) = Ax$ のクーブマン作用素であり、 $h(x) := Cx$ に対しては $(\mathcal{K}[h])(x) = CAx$ である。ここで、クーブマン作用素は任意の実数 λ および $h_1, h_2 \in L^\infty(\mathbb{X})$ に対して

$$(\mathcal{K}[\lambda h])(x) = \lambda \cdot h(f(x)) = \lambda (\mathcal{K}[h])(x), \quad (8.56)$$

$$(\mathcal{K}[h_1 + h_2])(x) = h_1(f(x)) + h_2(f(x)) = (\mathcal{K}[h_1] + \mathcal{K}[h_2])(x) \quad (8.57)$$

を満たすため線形作用素である。したがって、固有方程式 $\phi = \lambda \mathcal{K}[\phi]$ の解として与えられる固有値 $\lambda \in \mathbb{R}$ 、固有関数 $\phi \in L^\infty(\mathbb{X})$ が与えられれば、この作用素の性質が明らかになる。たとえば、

$$\phi(x_k) = (\mathcal{K}^k[\phi])(x_0) = \lambda^k \phi(x_0) \quad (8.58)$$

であるため、0に近い λ に対応する固有関数 ϕ で与えられる x_k の特徴量に注目しても、その値 $\phi(x_k)$ は時刻 k とともに直ちに0に収束する。言い換えると、絶対値が大きい λ に対応する固有関数 ϕ は、長期的に x_i の挙動を説明する特徴量であると言える。

一方で、クーブマン作用素の定義域は関数空間であり、与えられた非線形システム

$$x_{k+1} = f(x_k) \quad (8.59)$$

のクーブマン作用素 \mathcal{K} の固有値問題を厳密に解くことは現実的ではない。そこで、 f のクーブマン作用素 \mathcal{K} が与えられた関数の集合 $\{\phi_i\}_{i=1}^{N_f}$ で張られる部

分空間 $\text{span}\{\phi_i\} \subset L^\infty(\mathbb{R}^n)$ で閉じている, すなわち,

$$\mathcal{K}[\phi_i] = \sum_{j=1}^{N_f} a_{ij} \phi_j, \quad i \in \llbracket N_f \rrbracket \quad (8.60)$$

を満たす a_{ij} が存在するとしよう. この仮定のもとで固有値・固有関数を「 f を具体的に求めることなく」, x の時系列データから計算することを目的とする. ここで, $(\mathcal{K}[\phi_i])(x_k) = \phi_i(x_{k+1})$ であることに注意すると, $(A_f)_{ij} := a_{ij}$ の記法のもと (8.60) は (8.54) と等価である. したがって, データ $\{\phi(x_k)\}_k$ に動的モード分解を適用して A_f の主たる固有値と固有ベクトルを求めるることは, ($\text{span}\{\phi_i\} \subset L^\infty(\mathbb{R}^n)$ に限定して) 対応するクーブマンモードの固有値と固有関数の近似計算をおこなっていると解釈できる. また, 入力をもつシステムに対する拡張などもおこなわれている [32].

9 強化学習 (18)

強化学習あるいは強化という用語・概念は、もともとは生物の持つ適応能力・学習能力の研究の中で導入されたものである。その後、人工知能分野において人工物に同様の能力を持たせるための研究が行われるようになり、この用語が援用されるようになった。本章では、データ駆動型制御系設計に他ならない強化学習と、標準的なシステム制御の考え方の違いを整理した後、強化学習の代表的な手法を紹介する。

9.1 最適制御と強化学習

本章においても、基本的には問題 6.1.1 の無限区間最適制御問題を解くことが目標である。以下では、時刻 k におけるコストを確率過程 $\ell_k \in \text{rv}(\mathbb{R})$ によりあらわし、任意の k に対して

$$\varphi_{(x_{k+1}, \ell_k)}((x', l) | x_k = x, u_k = u) = \Psi_\ell(x', l | x, u) \quad (9.1)$$

が成り立つとする。これにより、 $\ell_k = \|x_k\|^2 + \|u_k\|^2 + \xi_k$, $\xi_k \in \text{rv}(\mathbb{R})$ といったコストの不確実性も表現できる。なお、

$$\ell(x, u) := \int_{\mathbb{R}} \int_{\mathbb{X}} l \cdot \Psi_\ell(x', l | x, u) dx' dl, \quad (9.2)$$

$$\Psi(x' | x, u) := \int_{\mathbb{R}} \Psi_\ell(x', l | x, u) dl \quad (9.3)$$

とおけば $\mathbb{E}[\ell_k] = \mathbb{E}[\ell(x_k, u_k)]$ であり、これまでの議論がそのまま成り立つ。したがって、最適制御則は状態フィードバックで与えられるため、

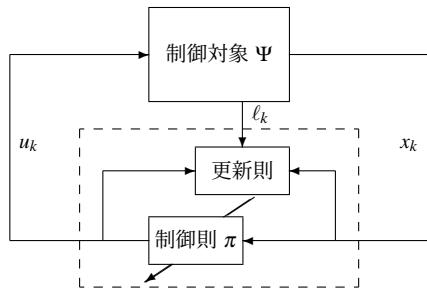


図 9.1: 強化学習制御器の構造

$$J(\pi) := \mathbb{E}^\pi \left[\sum_{k=0}^{\infty} \beta^k \ell_k \right] = \mathbb{E}^\pi \left[\sum_{k=0}^{\infty} \beta^k \ell(x_k, u_k) \right] \quad (9.4)$$

を最小化する状態フィードバック制御則 π^* を求める。これまでの章においては、モデルベースト設計と呼ばれる遷移確率度関数 Ψ_ℓ を利用可能であるという設定での解法（たとえば Ψ_ℓ で定まるベルマン方程式を解く）を紹介した。これに対して、 Ψ_ℓ は未知であるとし、 (x_k, u_k, ℓ_k) のデータを用いて制御則 π を更新する図 9.1 に示す構造の制御器を設計することを**強化学習** (reinforcement learning) と呼ぶ。主たる目的は制御則を π^* に収束させる更新則を設計することであるが、更新に有用な情報を得るために入力設計も更新則の役割である。

注意 9.1.1 (用語の対応)。システム制御における「制御対象・プラント／制御器・コントローラ／制御入力／制御則」という用語は、それぞれ「環境／エージェント／行動／方策・政策 (policy)」と置き換わる。また、システム制御では状態や入力が連続値をとる対象を扱うのに対して、強化学習分野では主に有限な状態・入力値をとる対象が主流であったが、近年では連続値をつかう強化学習の研究も増加している。 ◇

まず、8 章のシステム同定手法や 7 章の教師あり学習により、利用可能なデータからモデル Ψ_ℓ を求め、モデルベースト設計手法を適用することが考えられ

138 9. 強化学習 (18)

る。これとは対極的に、モデルを経由しない解法全般をさしてモデルフリー強化学習と呼ぶ。システム制御における直接型適応制御はモデルフリー設計に分類されるが、制御目的が最適制御よりも安定化や規範となるシステムの出力への追従であることが多い。一方で、データからモデルを構築し、モデルをシミュレータとしてデータをかさ増ししモデルフリー設計手法を用いること全般的にモデルベースト強化学習と呼ぶ。通常のシステム同定と比較すると、モデルの不確実性の学習や対処法にも重きをおいている場合が多い。間接型適応制御はモデルベーストに相当し、適応動的計画法 (adaptive dynamic programming) や近似動的計画法 (approximate ~) といった名称で研究されている。このような類義語については文献 [33] の §1.3.1 を確認していただきたい。モデルを構築することで、コスト関数のみが変化した場合にもそのまま適用できたり、対象に対する具体的な先見情報を組み込んだりと様々な利点がある。ただ単に強化学習手法といった場合はモデルフリー強化学習手法を指すことが多い。近年の目覚ましい発展のなかで、本節で紹介できるのは残念ながらそのごく一部である。すでに多くの書籍が出版されているが、システム制御研究者向けの追加の文献は [34] を参照してほしい。

9.2 モデルベースト設計

この節では、強化学習の考え方へ直結する価値関数とベルマン方程式の整理をしたのち、方策反復法・方策勾配法と呼ばれるモデルベースト手法を導入する。本節の段階では、データから適応的に制御則を獲得するという内容は含んでいないことに注意してほしい。

9.2.1 方策反復法

(1) ベルマン方程式 すでに状態価値関数は動的計画法の際に導入したが、類似の行動状態価値関数とあわせて (9.1) のもとで再定義する。状態フィードバック制御則 π に対して、

$$\mathcal{A}_\ell[V](x, u) := \int_{\mathbb{X}} \int_{\mathbb{R}} (l + \beta V(x')) \Psi_\ell(x', l|x, u) dl dx' \quad (9.5)$$

$$\mathcal{B}[V](x) := \inf_{u \in U(x)} \mathcal{A}_\ell[V](x, u) \quad (9.6)$$

$$\mathcal{B}^\pi[V](x) := \int_{U(x)} \mathcal{A}_\ell[V](x, u) \pi(u|x) du \quad (9.7)$$

- 状態価値関数

$$V^\pi(x) := \mathbb{E}_{x_0=x}^{\pi} \left[\sum_{k=0}^{\infty} \beta^k \ell_k \right] \quad (9.8)$$

$$V^*(x) := \inf_{\pi} V^\pi(x) \quad (9.9)$$

- 行動価値関数

$$Q^\pi(x, u) := \mathbb{E}_{x_0=x, u_0=u}^{\pi} \left[\sum_{k=0}^{\infty} \beta^k \ell_k \right] \quad (9.10)$$

$$Q^*(x, u) := \inf_{\pi} Q^\pi(x, u) \quad (9.11)$$

すでに得られているベルマン方程式を含めてまとめると,

$$Q^*(x, u) = \mathcal{A}_\ell[V^*](x, u) \quad (9.12)$$

$$V^*(x) = \mathcal{B}[V^*](x) = \inf_{u \in U(x)} Q^*(x, u) \quad (9.13)$$

$$Q^\pi(x, u) = \mathcal{A}_\ell[V^\pi](x, u) \quad (9.14)$$

$$V^\pi(x) = \mathcal{B}^\pi[V^\pi](x) = \int_{U(x)} Q^\pi(x, u) \pi(u|x) du \quad (9.15)$$

が成り立っている。動的計画法の骨子は入力系列を初期時刻とそれ以降に分割することであったが、各価値関数は表 9.1 にまとめた制御則により実現されるコストの値であることから、これらの方程式も自然に理解できるであろう。

すでに、最適制御性能に対応する V^* の計算法として定理 6.1.10 の価値反復法を与えた。ここでは制御則能 π に対する性能 V^π の計算が同様におこなえることを確認しておこう。

表 9.1: 状態価値関数・状態行動価値関数の入力

	u_0	$u_k, k \geq 1$
$V^*(x)$	$\pi^*(x)$	$\pi^*(x_k)$
$Q^*(x, u)$	u	$\pi^*(x_k)$
$V^\pi(x)$	$\pi(x)$	$\pi(x_k)$
$Q^\pi(x, u)$	u	$\pi(x_k)$

補題 9.2.1 (\mathcal{B}^π の単調性). 関数 $V_1, V_2 : \mathbb{X} \rightarrow \mathbb{R}$ が $V_1 \leq V_2$ を満たすならば,

$\mathcal{B}^\pi[V_1] \leq \mathcal{B}^\pi[V_2]$ である. \triangleleft

補題 9.2.2. 任意の関数 $V : \mathbb{X} \rightarrow \mathbb{R}$ および $i \geq 1$ に対して,

$$(\mathcal{B}^\pi)^i[V](x) = \mathbb{E}_{x_0=x}^{\pi} \left[\sum_{k=0}^{i-1} \beta^k \ell(x_k, u_k) + \beta^i V(x_i) \right] \quad (9.16)$$

が成り立つ. \triangleleft

この結果から、つぎの反復解法を示すことができる.

定理 9.2.3 (V^π に対する価値反復法). 問題 6.1.1において、 ℓ は有界な関数、 π を任意の状態フィードバック制御則とする。このとき、(9.8) の状態価値関数 V^π はベルマン方程式 (9.15) の唯一の有界な解であり、任意の上に有界な関数 $V : \mathbb{X} \rightarrow \mathbb{R}$ に対して,

$$\lim_{k \rightarrow \infty} \|(\mathcal{B}^\pi)^k[V] - V^\pi\|_\infty = 0 \quad (9.17)$$

が成り立つ. \triangleleft

証明. 定理 6.1.10 と同様に示される. \square

(2) 方策反復法 前節の方法などで状態価値関数 V^π が得られるとき、(9.14) により Q^π が求められる。つぎの定理は、 Q^π を用いた更新則

$$\pi'(x) := \arg \min_{u \in U(x)} Q^\pi(x, u) \quad (9.18)$$

により、制御則を改善していくことを示す。

定理 9.2.4 (方策改善定理 (policy improvement theorem)). 定理 9.2.3 と同じ仮定のもと,

$$Q^\pi(x, \pi'(x)) \leq V^\pi(x), \quad \forall x \in \mathbb{X} \quad (9.19)$$

を満たす任意の π' に対して,

$$V^{\pi'}(x) \leq V^\pi(x), \quad \forall x \in \mathbb{X} \quad (9.20)$$

が成り立つ. 逆に, V^π が有界かつ

$$Q^\pi(x, u) \geq V^\pi(x) - \varepsilon(1 - \beta), \quad \forall u \in U(x), \quad \forall x \in \mathbb{X}, \quad (9.21)$$

が成り立つとき,

$$V^\pi(x) \leq V^*(x) + \varepsilon, \quad \forall x \in \mathbb{X} \quad (9.22)$$

である. \triangleleft

証明. (9.19) は

$$(\ell + \beta \mathcal{A}[V^\pi])(x, \pi'(x)) \leq V^\pi(x) \quad (9.23)$$

と書けるため, (6.18) の導出と同様に (9.20) が示される. 後半の主張は,

$$(\ell + \beta \mathcal{A}[V^\pi - \varepsilon])(x, u) = Q^\pi(x, u) - \varepsilon\beta \geq V^\pi(x) - \varepsilon \quad (9.24)$$

が成り立つため, 注意 6.1.7 で $\check{V} := V^\pi - \varepsilon$ とすることで示される. \square

この定理の後半は, 初期時刻以外は π で固定したままで, 初期時刻の入力のみ変更して性能が改善しなければ, π は既に準最適である (初期時刻以外を同時に探索しても改善することはない) ことを示している. したがって, 方策反復法 (policy iteration) と呼ばれるアルゴリズム 4 の手順により最適制御問題を解くことができる.

Algorithm 4 方策反復法

Require: $\|\nabla^\pi\|_\infty$ が有限となる π で初期化

- 1: Q^π を求める.
- 2: $\pi'(x) \leftarrow \arg \min_{u \in U(x)} Q^\pi(x, u)$ とする.
- 3: $\|\pi' - \pi\|$ が十分小さければ π を出力して終了.
- 4: $\pi \leftarrow \pi'$ として 1. に戻る.

(3) LQR 問題に対する方策反復法 数種類の価値関数が登場し、現代制御理論とはまったく異なる理論体系に見えるかもしれないが、線形二次レギュレータ問題の場合

$$x_{k+1} = Ax_k + Bu_k + v_k, \quad l_k = \begin{bmatrix} x_k^\top & u_k^\top \end{bmatrix} \begin{bmatrix} Q & S \\ S^\top & R \end{bmatrix} \begin{bmatrix} x_k \\ u_k \end{bmatrix} \quad (9.25)$$

に限れば、やはり行列の計算で方策反復の手続きを記述することができる。制御則 π を $u_k = -Kx_k$ としたときの Q^π の計算の各ステップは以下の計算をすればよい。まず、

$$x_{k+1} = (A - BK)x_k + v_k, \quad \ell(x, u) = x^\top \begin{bmatrix} I & -K^\top \end{bmatrix} \begin{bmatrix} Q & S \\ S^\top & R \end{bmatrix} \begin{bmatrix} I \\ -K \end{bmatrix} x$$

に注意して、(9.8) の計算に定理 6.2.1 を適用すると、

$$\beta(A - BK)^\top \Pi(A - BK) + \begin{bmatrix} I & -K^\top \end{bmatrix} \begin{bmatrix} Q & S \\ S^\top & R \end{bmatrix} \begin{bmatrix} I \\ -K \end{bmatrix} = \Pi \quad (9.26)$$

の一意の正定解 Π および実定数 $\gamma > 0$ を用いて[†]、 $\nabla^\pi(x) = x^\top \Pi x + \gamma$ とあらわされる。また、(9.14) の $Q^\pi(x, u)$ の計算および(9.18) の最小化解 $u \in U(x)$ は補題 6.2.5 の証明と同様に計算できる。これらをまとめるとアルゴリズム 5 の手続きを得る。

価値反復法であるリッカチイタレーションでは、特別な初期化は不要であるが、反復中に得られる制御則は制御区間を有限で打ち切った問題 (6.25) の最適制御性能しか考慮しておらず、十分収束するまでは安定性も保証されなかった

[†] `dlyap(sqrt(beta)*(A-B*K)', [eye(n,n) -K']*[Q S'; S R]*[eye(n,n); K]);`

Algorithm 5 LQR 問題に対する方策反復法

Require: $\sqrt{\beta}\rho(A-BK) < 1$ を満たすゲイン K で初期化.

- 1: リアノフ方程式 (9.26) の一意正定解 Π を求める.
- 2: $K' \leftarrow K(\Pi) := (R + \beta B^\top \Pi B)^{-1} (S + \beta B^\top \Pi A)$ とする.
- 3: $\|K' - K\|$ が十分小さければ K を出力して終了.
- 4: $K \leftarrow K'$ として 1. に戻る.

(例 6.2.6). 一方で, 方策反復法は安定化制御則で初期化する必要があるが, 反復ごとに無限時間区間の制御性能が改善されることが保証されている.

例 9.2.5 (方策反復法). 例 6.2.6 の線形システムに方策反復法を適用する. ここで, $\Pi_{\text{ini}}^{\text{VI}}$ により初期化することで安定化制御則を与えた結果を図 6.1 に示している. 方策反復法では各反復においてリッカチ方程式を解く必要があるが, 誤差の収束は速い. \triangleleft

9.2.2 方策勾配法

本節では, 制御則をパラメータ $\mu \in \mathbb{R}^d$ で規定し, コスト関数の μ に関する勾配を用いて最適制御則を直接探す**方策勾配法** (policy gradient method) を説明する. 以下では, 確定的状態フィードバック $u_k = \pi_\mu(x_k)$ のみを扱うが, 確率的制御則に関しても同様の議論が可能である. なお, 導出において必要な微分・積分の順序交換可能性や極限操作の収束保証などは仮定しておく.

制御則 π_μ のもとでの評価関数 V^{π_μ} の μ に関する勾配を求めるにあたり, μ の変動に対してステージコストを与える関数 $x \mapsto \ell(x, \pi_\mu(x))$ がどう変化するかは単純に偏微分すればよいが, 代入すべき x_k の分布の変化は自明ではなく, つきの定理で特徴づけられる (証明は B.3).

定理 9.2.6 (確定方策に対する方策勾配定理). 確定方策 $\pi_\mu : \mathbb{X} \rightarrow \mathbb{U}$ のもとで,

$$\frac{\partial V^{\pi_\mu}}{\partial \mu}(x) = \mathbb{E}_{x_0=x}^{\pi_\mu} \left[\sum_{k=0}^{\infty} \beta^k \frac{\partial \pi_\mu}{\partial \mu}(x_k) \frac{\partial Q^{\pi_\mu}}{\partial u}(x_k, u_k) \right] \quad (9.27)$$

が成り立つ. \triangleleft

Algorithm 6 方策勾配法

Require: 制御則のパラメトリゼーション π_μ を与え, $\mu \in \mathbb{R}^d$ を初期化.

- 1: Q^{π_μ} を求める.
- 2: (9.28) により $\frac{\partial J(\pi_\mu)}{\partial \mu}$ を求め, $\mu' \leftarrow \mu - \alpha \frac{\partial J(\pi_\mu)}{\partial \mu}$ とする.
- 3: $\|\mu' - \mu\|$ が十分小さければ K を出力して終了.
- 4: $\mu \leftarrow \mu'$ として 1. に戻る.

制御則 π に対する問題 6.1.1 の評価関数値は $J(\pi) := \mathbb{E}[V^\pi(x_0)]$ である. 定理 6.1.5 で与えられる最適制御則 π^* は $J(\pi)$ のみならず任意の x に対する $V^\pi(x)$ を最小化する (定理 6.1.5 における初期分布の任意性). したがって, ある μ^* を用いて $\pi^* = \pi_{\mu^*}$ と書けるならば, $\mu = \mu^*$ および任意の x に対して (9.27) は 0 となる. しかしこれは, 任意の x に対する (9.27) を用いて, μ を更新すればよいということではない. 実際,

$$\frac{\partial J(\pi_\mu)}{\partial \mu} = \mathbb{E}_{x_0}^{\pi_\mu} \left[\frac{\partial V^{\pi_\mu}}{\partial \mu}(x_0) \right] = \mathbb{E}_{x_0}^{\pi_\mu} \left[\sum_{k=0}^{\infty} \beta^k \frac{\partial \pi_\mu}{\partial \mu}(x_k) \frac{\partial Q^{\pi_\mu}}{\partial u}(x_k, u_k) \right] \quad (9.28)$$

を計算して, $J(\pi_\mu)$ の減少方向を求める必要がある (アルゴリズム 6).

注意 9.2.7. 方策反復法において, (9.18) の π' を求めるには,

$$\mathcal{C} := \{(x, u) : x \in \mathbb{X}, u \in U(x)\} \quad (9.29)$$

のすべての点に対するに対する関数値 $Q^\pi(x, u)$ が必要である. これに対して, 方策勾配法において必要となるのは, $x \in \mathbb{X}$ に対する

$$\left. \frac{\partial Q^{\pi_\mu}(x, u)}{\partial u} \right|_{u=\pi_\mu(x)} \quad (9.30)$$

の値のみである. したがって, 各 x に対して $U(x)$ 全体ではなく $\pi(x)$ の近傍の u に対するよい近似値が得られていれば十分である (図 9.2). これは制御則を微小に変化させたときに制御性能の変化を求める式であることを考えると, 妥当な情報量であるといえる. \diamond

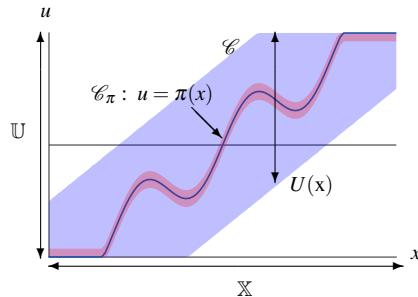


図 9.2: 状態行動価値関数を求める領域

9.3 モデルフリー設計

本節では、(9.1) の Ψ_ℓ が未知である状況での強化学習手法として、モデルを用いずに方策反復法や方策勾配法を実装する方法を考察する。

9.3.1 行動価値関数の推定

(1) TD 誤差学習 方策反復法、方策勾配法のどちらにおいても、モデルを用いていたのは与えられた制御則 π に対して行動価値関数 Q^π を求める箇所のみであった。そこで、制御対象の入出力データから、 Q^π を関数回帰するつぎの問題を考える。

問題 9.3.1 (状態行動価値関数の推定). (9.1) の確率密度関数 Ψ_ℓ および状態フィードバック制御則 π に対して、 x_k, u_k, ℓ_k の標本系列 $(x_k, u_k, l_k)_{k \in \mathbb{Z}_+}$ を用いて、行動価値関数 Q^π を推定せよ。 \diamond

この問題に対して、いくつかの点における Q^π の代入値（の近似）が得られれば通常の関数回帰手法を用いることができるが、データ $(x_k, u_k, l_k)_{k \in \mathbb{Z}_+}$ から直接的にこのような情報を得ることは難しい。そこで、TD 誤差 (temporal difference error) と呼ばれる $\ell(x_k, u_k) + \beta Q^\pi(x_{k+1}, \pi(x_{k+1}))$ を $Q^\pi(x_k, u_k)$ の近似とした場合の誤差に対応する

$$\delta_k := \ell_k + \beta \hat{Q}^\pi(x_{k+1}, \pi(x_{k+1})) - \hat{Q}^\pi(x_k, u_k) \quad (9.31)$$

を定義する。このとき、行動価値関数に関するベルマン方程式 (9.14) より、
 $\hat{Q}^\pi = Q^\pi$ ならば

$$\mathbb{E}[\delta_k | x_{:k}, u_{:k}, \ell_{:k-1}] = 0, \forall k, \mathbf{x}_{:k}, \mathbf{u}_{:k-1}, \mathbf{l}_{:k-1} \quad (9.32)$$

が成り立つ。これに注目して、以下では δ_k が 0 に近い値をとるような関数 \hat{Q}^π を Q^π の推定結果とするという立場をとる。たとえば、各組 (x_k, u_k, l_k, x_{k+1}) に対する誤差の和

$$\sum_k \|l_k + \beta \hat{Q}^\pi(x_{k+1}, \pi(x_{k+1})) - \hat{Q}^\pi(x_k, u_k)\|^2 \quad (9.33)$$

の最小化をおこなえばよいが、 $\pi(x_{k+1})$ を u_{k+1} としたり、 u_k を $\pi(x_k)$ としないように注意されたい。

方策反復法と方策勾配法のどちらにおいても、暫定的に最善の制御則 π に対する Q^π を求める必要がある。したがって、性能の観点からも問題 9.3.1 を解きたい状況では、 $u_k = \pi(x_k)$ のもとの閉ループデータが与えられることが多いと考えられる。ここで入力データに関連して、いくつかの注意を述べる。

- (9.33) を最小化する際のデータは、 $u_k = \pi(x_k)$ のもとの閉ループデータである必要はない。したがって、過去に π とは無関係に取得・蓄積されたデータも用いることができる。
- 確定的な状態フィードバック $u_k = \pi(x_k)$ に従っていると、行動価値関数の推定がその定義域のごく一部である図 9.2 の \mathcal{C}_π 上でしかおこなわれない。とくに x_k, u_k がそれぞれ \mathbb{X}, \mathbb{U} 全体をわたっているだけでは不十分であることに注意する。これについては、例えば
 - 雑音 z_k を付加して

$$u_k = \pi(x_k) + z_k, \quad (9.34)$$

- 小さな $\varepsilon > 0$ に対して、確率 $(1 - \varepsilon)$ で $\pi(x_k)$ 、確率 ε で一様乱数とする (ε -greedy と呼ばれる)

などにより、現在の制御則 π によって計算される（性能が良いと考えられる）制御入力以外もあえて使われるようにして解決が図られる。こうした性能と情報取得のトレードオフは「探索（Explore）と利用（Exploitation）」のトレードオフ」と呼ばれる。システム同定の知識があれば、PE 性の確保と結び付けて理解できるであろう。

また、ある制御則 π に対する Q^π の推定に要するデータ量および計算量にも注意が必要である。オフラインで最適制御則を導くのであれば、十分な精度で推定値が求まる（収束する）まで計算を続けてから、制御則の改善を行うのが良いと考えられる。しかしながら、大規模な問題となるとこの収束に時間がかかるだけでなく、強化学習問題では推定値の更新のために制御対象に実際に制御入力を印加しなければならないため、最適でない制御則にもとづいて制御がおこなわれることはある種の損失を生んでいることになる。そこで、実用的には十分な精度で推定値が求まる前に制御則が更新される。

最も極端な場合には (9.33) 式の演算を一組のデータのみを用いた推定値にもとづいて制御則更新が行われる。たとえば、 \mathbb{X}, \mathbb{U} が有限集合であるマルコフ決定過程の場合、特定の点における \hat{Q}^π の関数値を容易に更新できるため、

$$\hat{Q}^\pi(x_k, u_k) \leftarrow \hat{Q}^\pi(x_k, u_k) + \alpha_k(l_k + \beta \hat{Q}^\pi(x_{k+1}, \pi(x_{k+1})) - \hat{Q}^\pi(x_k, u_k)) \quad (9.35)$$

と更新する場合もある。この $\alpha_k \in [0, 1]$ は学習係数（学習率）あるいはステップサイズと呼ばれる更新の大きさを調整する実数値である。が、このようなケースでも、適切な条件の下では最適制御則への収束を示すことができる [35][†]。

注意 9.3.2. TD 誤差は、場合によって定義式に差異はあるものの、強化学習手法において頻出する。たとえば、 Q^π ではなく Q^* を推定する場合、(9.12) より、

$$\delta_k^* := l_k + \beta \inf_{u \in U(x_{k+1})} \hat{Q}^*(x_{k+1}, u) - \hat{Q}^*(x_k, u_k) \quad (9.36)$$

および

[†] SARSA の詳細と、確率近似法に基づいた厳密な議論については、例えば文献 [36] を参考いただきたい。Q 学習

$$\widehat{Q}^*(x_k, u_k) \leftarrow \widehat{Q}^*(x_k, u_k) + \alpha_k(l_k + \beta \inf_{u \in U(x_{k+1})} \widehat{Q}^*(x_{k+1}, u) - \widehat{Q}^*(x_k, u_k)) \quad (9.37)$$

となる。 ◇

(2) RLS アルゴリズムの利用 ところで、 \mathbb{X}, \mathbb{U} が連続的な値をとる集合の場合は、なんらかの正則化もしくは関数形の限定が必要であることは、7.1.1 節で述べた通りである。ここでは、特徴写像 $\phi: \mathbb{X} \times \mathbb{U} \rightarrow \mathbb{R}^{N_f}$ を用いて、

$$\widehat{Q}^\pi(x, u) := \phi(x, u)^\top \theta_\pi, \quad \theta_\pi \in \mathbb{R}^{N_f} \quad (9.38)$$

とかける関数によって行動価値関数 Q^π を推定する。つまり、(9.33) の

$$\sum_k \|l_k - \{\phi(x_k, u_k) - \beta \phi(x_{k+1}, \pi(x_{k+1}))\}^\top \theta_\pi\|^2 \quad (9.39)$$

を最小化する θ_π を点推定値として用いる。これは適応フィルタで考えた問題そのものであるため、8.1.1 節の RLS フィルタにより逐次更新できる。すなわち、 θ_π の k 番目の推定値 $\hat{\theta}_\pi(k)$ を

$$\begin{aligned} \phi_k &:= \phi(x_k, u_k) - \beta \phi(x_{k+1}, \pi(x_{k+1})) \\ \hat{\theta}_\pi(k) &:= \hat{\theta}_\pi(k-1) + \frac{P_\pi(k-1) \phi_k (l_k - \phi_k^\top \hat{\theta}_\pi(k-1))}{1 + \phi_k^\top P_\pi(k-1) \phi_k} \\ P_\pi(k) &:= P_\pi(k-1) - \frac{P_\pi(k-1) \phi_k \phi_k^\top P_\pi(k-1)}{1 + \phi_k^\top P_\pi(k-1) \phi_k} \end{aligned}$$

のように構成できる。

特徴関数の取り方に関しては、制御対象が線形システムかつ π が安定化線形フィードバック制御則のとき、 Q^π は $[x; u]$ の二次形式であるという情報が活用できる。具体的には、係数行列 Υ^π の要素と γ_π を並べたベクトル θ_π (Υ^π の対称性から $(n_x + n_u)(n_x + u_u + 1)/2 + 1$ 次元)、 x, u の要素の二次単項式および定数 1 を並べたベクトル $\phi(x, u)$ を用いて、

$$Q^\pi(x, u) = \begin{bmatrix} x^\top & u^\top \end{bmatrix} \Upsilon^\pi \begin{bmatrix} x \\ u \end{bmatrix} + \gamma_\pi = \phi(x, u)^\top \theta_\pi \quad (9.40)$$

と過不足なくあらわせる。

9.3.2 アルゴリズム例

ここでは、前節の TD 誤差学習で状態行動価値関数が得られることを前提として、制御則の更新方法の具体例を紹介する。

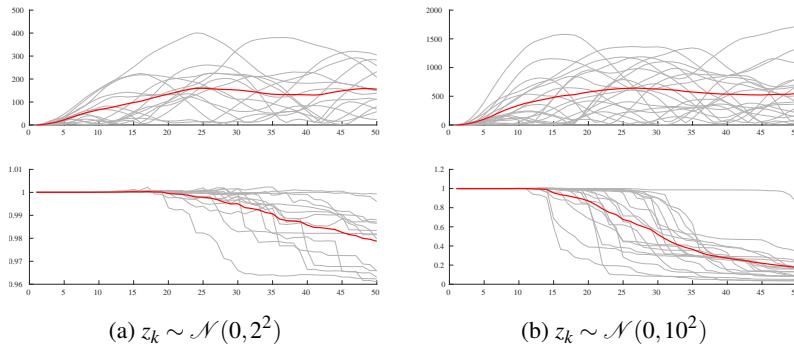
(1) 方策反復法 離散値をとるシステムの TD 誤差学習を方策反復法に組み込んでみよう。具体的には、(9.35)による推定行動価値関数の更新、(9.18)による制御則（必要に応じて ϵ -greedy）の更新を毎時刻おこなう手法を SARSA 法[†]と呼ぶ。同様に、行動価値関数の推定を (9.37) におきかえた手法を Q 学習と呼ぶ。いずれも制御則を毎時刻更新するため π' の関数形は不要で $\pi'(x_{k+1})$ の値のみ計算すればよい、網羅的な評価により求められる。ただし、 u_k の決定に暫定的な制御則 π を用いない場合、SARSA 法の (9.35) の右辺の計算には π への代入が必要（方策オン型という）であるのに対して、Q 学習における更新式には π があらわれない（方策オフ型という）という違いがある。

一方で、連続値をとるシステムの場合、 \mathcal{C} 全体で Q^π を求めるのも、(9.18) を計算することも容易ではない。幸い、これまでに説明したとおり線形システムの場合であればどちらも効率的に実行可能である。実際、 (A, B) が可制御であれば、このアルゴリズム 4 は、最適なフィードバックゲインに収束することが証明できる [37]。

例 9.3.3 (TD 学習と方策反復法にもとづく強化学習). 例 6.2.6 の線形システムに対して、初期安定化制御則 $\pi_0(x) = -K_0x := [4.1100, 11.7519, 19.2184]x$ を用いて ($\rho(A - BK_0) = 0.88$)、TD 学習をおこなった結果を図 9.3 に示す。具体的には、 $u_k = -K_0x_k + \xi_k$ で制御された際の軌道 $\{(x_k, u_k, \ell_k)\}_k$ の観測値にもとづく RLS アルゴリズムにより Q^{π_0} の推定をおこない、毎時刻更新される行列を \hat{Y}_k と表記する。また、細線は 20 本の標本経路、太線はその平均値である。

まず、 $\xi_k \sim \mathcal{N}(0, 2^2)$ の場合は、状態軌道の変動は大きくないものの、学習はまったく進んでいない。実際、ほぼすべての標本経路において \hat{Y}_{50} からは安定化ゲインは得られなかった。一方で、 $\xi_k \sim \mathcal{N}(0, 10^2)$ の場合は、状態軌道の変

[†] あ

図 9.3: TD 学習 (上段: 状態変動 $\|x_k\|$, 下段: 相対近似誤差 $\|\hat{Y}_k - Y^{\pi_0}\| / \|Y^{\pi_0}\|$)

動が大きいものの、誤差は小さくなっている。

続いて, $\xi_k \sim N(0, 10^2)$ として, RLS アルゴリズム 50 ステップごとに制御則を更新した結果を図 9.4 に示す。ただし, 更新ごとに $\hat{\Theta}_{50i} = O$, $P_{50i} = 10^4 \times I$, $i = 0, 1, 2, \dots$ と初期化した。適応的に得られる制御則は最終的に最適制御則に収束し、状態変数の変動も小さく抑えられていることがわかる。 ◇

(2) 方策勾配法 つぎに、方策勾配法の実装を考える。幸い、(9.28) はモデルを用いることなく計算するのに適した形となっている。すなわち、制御則 π_μ に従って、 N_f ステップの制御を M エピソード行って得られた標本 $(x_k^m, u_k^m)_{k=0, \dots, N_f-1}^{m=0, \dots, M-1}$ が得られたとき、

$$\frac{\partial J(\pi_\mu)}{\partial \mu} \simeq \frac{1}{M} \sum_{m=0}^{M-1} \sum_{k=0}^{N_f-1} \beta^k \frac{\partial \pi_\mu}{\partial \mu}(x_k^m) \frac{\partial \hat{Q}^{\pi_\mu}}{\partial u}(x_k^m, u_k^m) \quad (9.41)$$

のように期待値を標本近似できる。制御則 π_μ は設計者が決定できる関数であるため、その勾配も既知である。また、行動状態価値関数 Q^{π_μ} の近似 \hat{Q}^{π_μ} が TD 誤差学習により解析的に得られていれば、その勾配も求められる。このように制御則評価の役割を担う行動状態価値関数 (critic とも呼ばれる) をもとに、(9.41) から標本近似で勾配を計算し、勾配法により制御則 (actor とも呼ばれる) を改善する。更新した制御則に従って状態値などの標本を取得し、改めて行動状態価値関数を求める。これを繰り返すことで最適な制御則を学習する。このように制

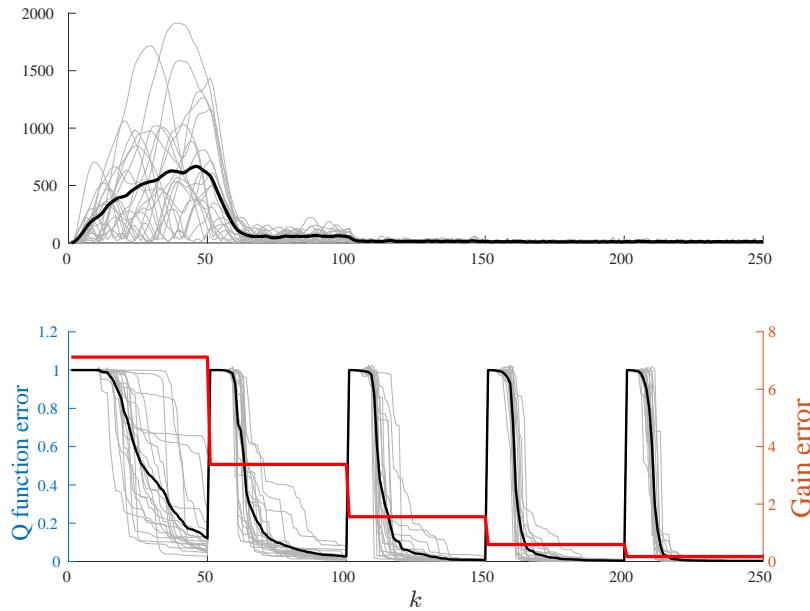


図 9.4: 強化学習（上段：状態変動 $\|x_k\|$ ，下段：相対近似誤差 $\|\hat{Y}_k - Y^{\pi_i}\|/\|Y^{\pi_i}\|, \|\hat{K}_i - K^*\|/\|K^*\|\right)$

御則である actor とそれを評価する critic の 2 モジュールで学習する枠組みは一般に actor-critic 法と呼ばれる。特に， $\hat{Q}^{\pi_\mu} : \mathbb{X} \times \mathbb{U} \rightarrow \mathbb{R}$ や $\pi_\mu : \mathbb{X} \rightarrow \mathbb{U}$ をニューラルネットワークで表現する DDPG (Deep Deterministic Policy Gradient) と呼ばれる手法では，アルゴリズム中にあらわれる勾配計算が自動微分によって効率的に計算できる。ただし，学習を安定化するためには，多少乱暴ではあるが慣例的に割引率を無視して，

$$\frac{\partial J(\pi_\mu)}{\partial \mu} \simeq \frac{1}{M} \sum_{m=0}^{M-1} \frac{1}{T} \sum_{k=0}^{T-1} \frac{\partial \pi_\mu}{\partial \mu}(x_k^m) \frac{\partial \hat{Q}^{\pi_\mu}}{\partial u}(x_k^m, u_k^m) \quad (9.42)$$

のように近似したり，ターゲットネットワークと呼ばれる仕組みを用いるなど，様々な工夫が必要となる。こうした枠組みもすでに既存のパッケージに標準的に実装されているので，上記の大枠がつかめたらまずは実際に試してみることをお勧めしたい。

9.3.3 部分観測問題

ここまででは完全観測制御問題、すなわち $u_k \in \sigma(x_k)$ のもとで強化学習手法をいくつか導出してきた。つぎに $u_k \in \sigma(y_{0:k-1})$ の部分観測制御問題を考える。現状では、部分観測制御問題に対して、確立したモデルフリー制御手法はそれほど知られていない。ここでは、線形確定システムに限定した問題 5.2.3 のみを紹介する。

システム同定の章で考察したように、状態変数は過去の出力及び入力系列から構成する必要があるが、どの程度の長さの系列が必要かは決定することは困難である。ただし、線形システムが対象である場合、可観測性の仮定のもと、必要な系列長を決定することができる。

定理 9.3.4. (5.3) の記法のもと、自然数 \bar{n} が $\text{rank } O_{\bar{n}} = n$ を満たすとき、

$$x_k = \check{T} \check{x}_k := \check{T} \underbrace{\begin{bmatrix} y_{k-\bar{n}}; \dots; y_{k-1}; u_{k-\bar{n}}; \dots; u_{k-1} \end{bmatrix}}_{=:\check{x}_k} \quad (9.43)$$

を満たす \check{T} が存在する。 \triangleleft

証明. まず、 $M_j := [A^j B, \dots, AB, B]$, $\bar{B}_0 := M_{\bar{n}-1}$, $\bar{B}_i := [M_{\bar{n}-i-1}, O]$, $\bar{B}_{\bar{n}} := O$ とおくと、

$$x_{k-i} = A^{\bar{n}-i} x_{k-\bar{n}} + \bar{B}_i \begin{bmatrix} u_{k-\bar{n}}; \dots; u_{k-1} \end{bmatrix}, \quad i = 0, \dots, \bar{n} \quad (9.44)$$

が成り立つ。これに左から C をかけて $i = \bar{n}, \dots, 1$ の順に縦に並べると、

$$O_{\bar{n}} x_{k-\bar{n}} = \begin{bmatrix} I & -[C \bar{B}_{\bar{n}}; \dots; C \bar{B}_1] \end{bmatrix} \check{x}_k \quad (9.45)$$

とかける。したがって、フルランク性より存在性が保証される擬似逆行列 $O_{\bar{n}}^\dagger$ を左からかけて、 $i = 0$ とおいた (9.44) に代入すると求める結果がしたがう。 \square

この結果と \check{x}_k の定義より、 $u_:$ と $y_:$ の関係は適切な \check{A} , \check{B} , \check{C} を用いて

$$\check{x}_{k+1} = \check{A} \check{x}_k + \check{B} u_k, \quad y_k = \check{C} \check{x}_k \quad (9.46)$$

と表される。また、ステージコストは x_k, u_k の二次形式なので、 \ddot{x}_k, u_k の二次形式で表現できる。したがって、 $\text{rank } O_{\bar{n}} = n$ を満たす \bar{n} が既知でさえあれば、状態変数 \ddot{x}_k が直接観測できる線形システムの LQR 問題として、これまでのモデルフリー手法がそのまま適用でき、ある行列 \check{K} を用いて $u_k = -\check{K}\ddot{x}_k$ とあらわされる最適制御則が求められる。ここで、必要なのは定数 \bar{n} のみであり、行列 $\check{T}, \check{A}, \check{B}, \check{C}$ などは必要はないことに注意する。

文 献 案 内

- 最近の書籍 [33, 38–41] <<<< HEAD =====
- Algorithms for Reinforcement Learning [36]
- これからの強化学習 [42]
- 脳神経科学 [43–48]

表??では制御対象が動的システムである（制御対象に状態が存在する）ことを前提としたが、もちろん、制御対象が静的システムの場合も考えられる。この場合は、ブラックボックス最適化問題や多腕バンディット問題の名で呼ばれる。

10**マルコフ連鎖と平衡状態(19)**

ここまでこの章では、システム制御理論で標準的である状態変数や入力変数が連続値をとる確率システムを想定した議論をおこなってきた。本章では、状態空間や入力空間が有限集合である場合に特化して、確率システムの表現や最適制御問題の解法を振り返る。また、確率システムの場合には自明ではない平衡状態について検討する。

10.1 有限状態確率システム**10.1.1 ベルマン方程式と線形計画法**

(1) マルコフ決定過程 状態変数と入力変数が有限個の値しかとらない有限状態確率システムを対象とした確率制御問題は**マルコフ決定過程** (Markov Decision Process; MDP) と呼ばれ、機械学習分野などで幅広く研究がおこなわれてきた。第4章の定理4.1.5や定理4.1.7などの動的計画法に関する結果は、密度関数を確率に置き換えるだけで直接的に適用できる。具体的には、状態変数が $\llbracket n \rrbracket := \{1, 2, \dots, n\}$ 、入力変数が $\llbracket m \rrbracket := \{1, 2, \dots, m\}$ にそれぞれ値をとる確率システムの遷移確率が

$$\mathbb{P}(x_{k+1} = x' | x_k = x, u_k = u) = \Psi_k(x' | x, u) \quad (10.1)$$

で与えられるとする。このとき、(4.29)に対応して

$$\mathcal{A}_x[V](k, x, u) := \sum_{x' \in \llbracket n \rrbracket} \{V(k+1, x') \Psi_k(x' | x, u)\} \quad (10.2)$$

と表記すると、定理 4.1.7 の 最適性原理（状態価値関数がベルマン方程式を満たす）はそのまま成立する。同様に、

$$\mathcal{A}_x[V](x, u) := \sum_{x' \in \llbracket n \rrbracket} \{V(x')\Psi(x'|x, u)\} \quad (10.3)$$

と表記すると、無限区間の場合の定理 6.1.4 ($V^* = \mathcal{B}[V^*]$) も成り立つ。

注意 10.1.1. 確率密度関数から確率への置き換えには、状態空間を \mathbb{R} とし、
 $\varphi_{x_k}(\cdot) = \sum_{x \in \llbracket n \rrbracket} \mathbb{P}(x_k = x) \delta_x(\cdot)$ などとみなしてもよい。ただし、 n 個の各状態に便宜上 $x \in \llbracket n \rrbracket$ によりラベル付けただけであり、整数の構造（状態 1 は状態 5 より状態 2 に近いなど）を取り入れたわけではない。こうした変数を**カテゴリー変数**と呼び、最適化などにおいて意図せずこうした構造を入れてしまわないようにワンホット（One-hot）ベクトルを用いるなどの注意を要する。 ◇

マルコフ決定過程の場合、ベルマン方程式に含まれる最小化はたかだか $|U|$ 個の値の比較であるため、状態価値関数は網羅的な有限回の評価で計算可能であり、最適制御則も同様に構成できる。これに対して、以下では少し視点を変えて、連続最適化により状態価値関数の上下限を求ることを考えてみよう。

(2) 下限 系 4.2.2 を用いると、マルコフ決定過程の場合の問題 4.1.4 の最適制御性能の下限は、

$$\underset{\{\check{V}(k,x)\}_{k,x}}{\text{maximize}} \quad \sum_{x \in \llbracket n \rrbracket} \mathbb{P}(x_0 = x) \check{V}(0, x) \quad (10.4)$$

$$\text{subj. to} \quad \check{V}(k, x) \leq \ell(k, x, u) + \sum_{x' \in \llbracket n \rrbracket} \Psi_k(x'|x, u) \check{V}(k+1, x') \quad (10.5)$$

$$\check{V}(\bar{k}, x) \leq \ell_f(x) \quad (10.6)$$

を解くことでのが求められる。なぜなら、この不等式制約を満たす任意の \check{V} に対して $V^* \geq \check{V}$ が成立し、 $\mathbb{E}_{x_0}[\check{V}(0, x_0)] = \sum_{x \in \llbracket n \rrbracket} \mathbb{P}(x_0 = x) \check{V}(0, x)$ が成り立つからである。また、この問題は決定変数 $\{\check{V}(k, x)\}_{k,x}$ に関する線形計画問題であり、ある程度効率的に解くことができる。無限区間においても、系 6.1.6 より、

156 10. マルコフ連鎖と平衡状態 (19)

$$\underset{\{\check{V}(x)\}_x}{\text{maximize}} \quad \sum_{x \in \mathbb{I}^n} \mathbb{P}(x_0 = x) \check{V}(x) \quad (10.7)$$

$$\text{subj. to} \quad \check{V}(x) \leq \ell(x, u) + \beta \sum_{x' \in \mathbb{I}^n} \Psi(x'|x, u) \check{V}(x), \forall x, u \quad (10.8)$$

を解くことで最適制御性能の下限が求められる。なお、 \check{V} は上に有界な関数であるため (6.17) はいつでも成立する。

(3) 上限 任意に与えられた確率的状態フィードバック π に対して定まる

$$\zeta_k(x, u) := \mathbb{P}^\pi(x_k = x, u_k = u) \quad (10.9)$$

を定義する。明らかに、

$$\left. \begin{array}{l} \zeta_k(x, u) \geq 0, \forall k, x, u \\ \sum_x \sum_{u \in U_k(x)} \zeta_k(x, u) = 1, \forall k \\ \sum_{u' \in U_{k+1}(x')} \zeta_{k+1}(x', u') = \sum_x \sum_{u \in U_k(x)} \Psi_k(x'|x, u) \zeta_k(x, u), \forall k \\ \sum_{u \in U_0(x)} \zeta_0(x, u) = \mathbb{P}(x_0 = x) \end{array} \right\} \quad (10.10)$$

である。逆に、この条件を満たす任意の $\{\zeta_k(x, u)\}_{k,x,u}$ に対して、(10.9) を満たす確率的状態フィードバック $\pi_k(u|x)$ が構成できる（定理 4.1.5 の証明と同様）。したがって、

$$\underset{\{\zeta_k(x, u)\}_{k,x,u}}{\text{minimize}} \sum_k \ell(k, x, u) \zeta_k(x, u) \quad \text{subj. to} \quad (10.10)$$

を解くことで最適制御性能の上限が求まる。

無限区間に關しては、

$$\zeta(x, u) := \sum_{k=0}^{\infty} \beta^k \mathbb{P}^\pi(x_k = x, u_k = u) \quad (10.12)$$

を定義すると、拘束条件

$$\left. \begin{aligned} \zeta(x, u) &\geq 0, \forall x, u \\ \sum_{u \in U(x)} \zeta(x, u) &= \mathbb{P}(x_0 = x) + \beta \sum_{x', u'} \Psi(x|x', u') \zeta(x', u'), \quad \forall x \end{aligned} \right\} \quad (10.13)$$

を得る。したがって、

$$\underset{\{\zeta(x, u)\}_{x, u}}{\text{minimize}} \sum_{x, u} \ell(x, u) \zeta(x, u) \quad \text{subj. to (10.13)} \quad (10.14)$$

を解くことで最適制御性能の上限が求まる。なお、下限のために導入した線形計画問題とは、数理最適化の意味での主双対関係にある。

演習 10.1. (10.13) を導出せよ. △

10.1.2 線形可解マルコフ決定過程

本節の目標は、マルコフ過程を定理 3.1.7 のような確率分布の遷移として表現することに慣れることと、カルバック・ライブラー制御と呼ばれる最大エントロピー制御に対応する制御問題の構造を調べることである。なお、ベクトル p の i 番目の要素を $(p)_i$ 、行列 P の (i, j) 成分を $(P)_{ij}$ 、 j 列を $P_{:,j}$ と表記する。

定義 10.1.2 (マルコフ連鎖). 整数 $n > 0$ に対して、非負行列 $P \in \mathbb{R}^{n \times n}$ が確率行列 (stochastic matrix)，すなわち

$$P^\top \mathbb{1} = \mathbb{1} \quad (10.15)$$

を満たすとする。このとき、

$$\mathbb{P}(x_{k+1} = i | x_k = j) = (P)_{ij}, \quad \forall i, j \in \llbracket n \rrbracket \quad (10.16)$$

を満たす $\llbracket n \rrbracket$ -値マルコフ過程 x を遷移行列 P をもつマルコフ連鎖 (Markov chain) と呼ぶ。 △

この確率行列は遷移確率密度関数に対応するが、状態変数の確率密度関数に対応するベクトルを導入すると、つぎの簡単な表現が得られる。

定理 10.1.3 (マルコフ連鎖の確率分布の時間発展). 遷移行列 P をもつマルコフ連鎖 x_i に対して, $(p_k)_i := \mathbb{P}(x_k = i)$ により定義される非負ベクトル $p_k \in \mathbb{R}^n$ は確率ベクトル (stochastic vector), すなわち

$$(p_k)_i \geq 0, \mathbb{1}^\top p_k = 1 \quad (10.17)$$

を満たし,

$$p_{k+1} = P p_k, \forall k \in \mathbb{Z}_+ \quad (10.18)$$

が成り立つ. \triangleleft

(1) カルバック・ライブラーア制御 以下では, 定理 10.1.3 の表現を用いて, 特殊な入力コストを有するつぎの制御問題を考える.

問題 10.1.4 (カルバック・ライブラーア制御). 遷移行列 P^π をもつマルコフ過程 p_k^π を考える. このとき, 初期分布 p_0 , 状態ステージコスト $\ell \in \mathbb{R}^n$, 割引率 $\beta \in (0, 1)$, 遷移行列 P^0 に対して,

$$\sum_{k=0}^{\infty} \beta^k \left\{ \ell^\top p_k^\pi + D_{\text{KL}}(P^\pi p_k^\pi \| P^0 p_k^\pi) \right\} \quad (10.19)$$

を最小化する遷移行列 P^π を求めよ. ただし, 確率ベクトル p, q に対して,

$$D_{\text{KL}}(p \| q) := \sum_{i=1}^n (p)_i \log \frac{(p)_i}{(q)_i} \quad (10.20)$$

と定義する. \triangleleft

ここで, $(\ell)_i$ は状態 i でのコストをあらわす. 一方, 制御しない場合の遷移確率が P^0 であり, 遷移行列を P^0 から P^π に整形することが制御による効果である. この整形にかかるコストを遷移後の状態分布のカルバック・ライブラーア距離で定量化している. この距離が有界となるためには, $(P^0 p_k)_i = 0$ ならば $(P^\pi p_k)_i = 0$ でなければならない. これは制御しない場合に可能な状態遷移しか許容しないことに対応する. 逆に, この制約さえ満たしていればいかなる遷移も許容する.

この問題を解くにあたり、状態価値関数を用意する必要がある。状態が確定的に $x_k = j$ となることは、確率ベクトル p_k が第 j 成分のみが 1、他は 0 の基本ベクトル e_j であることに対応する。そこで、初期状態 $p_0 = e_j$ の場合の (10.19) の下界値を第 j 要素とするベクトル $V^* \in \mathbb{R}^n$ が状態価値関数である。また、 $\Psi(i|j, u)$ に対応するのは P^π の (i, j) -成分であり、(10.3) の $\mathcal{A}_x[V](j, u)$ に対応するのは、 $V^\top P^\pi e_j$ である。これらをまとめると、つぎの結果を得る。

定理 10.1.5 (KL 制御問題の解)。 問題 10.1.4 に対して、最適制御則 P^π は

$$(P^*)_{ij} := \frac{(Z^*)_i (P^0)_{ij}}{Z^{*\top} (P^0)_{:,j}}, \quad (10.21)$$

で与えられる。ここで状態価値関数 V^* に対して

$$(Z^*)_j := \exp(-\beta(V^*)_j) \quad (10.22)$$

により定義される適合関数 Z^* は

$$(Z)_j^{1/\beta} = \exp(-\ell_j) (Z^\top P^0)_j, \quad \forall j \in \llbracket n \rrbracket \quad (10.23)$$

の解である。 ◇

証明. ベルマン方程式 (6.8) に $x = j$ ($p = e_j$) を代入すると、 V^* は方程式

$$(V^*)_j = \ell_j + \min_\pi \left\{ D_{\text{KL}}((P^\pi)_{:,j} \| (P^0)_{:,j}) + \beta (V^*)^\top (P^\pi)_{:,j} \right\} \quad (10.24)$$

を満たす。ここで、最小化すべき関数は、

$$\sum_{i=1}^n (P^\pi)_{ij} \left\{ \log \frac{(P^\pi)_{ij}}{(P^0)_{ij}} + \beta (V^*)_i \right\} \quad (10.25)$$

$$= \sum_{i=1}^n (P^\pi)_{ij} \left\{ \log \frac{(P^\pi)_{ij}}{(P^0)_{ij} (Z^*)_i} \right\} \quad (10.26)$$

$$= D_{\text{KL}}((P^\pi)_{:,j} \| (P^0)_{:,j}) - \log((Z^*)^\top P^0)_j \quad (10.27)$$

である。さらに (10.27) の第二項は P^π に依存しないことに注意すると、 $P^\pi = P^*$ が最小化解であり、最小値は $-\log((Z^*)^\top P^0)_j$ である。以上の議論より、 Z^* は

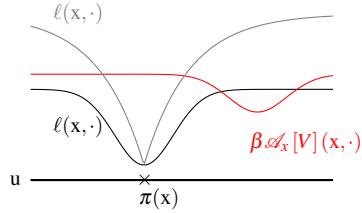
$$(V^*)_j = \ell_j - \log(Z^{*\top} P^0)_j = -(1/\beta) \log(Z^*)_j \quad (10.28)$$

を満足するので、(10.23) がしたがう。 \square

ここで導入した関数 Z^* を求める (10.23) は、 $\beta = 1$ のとき線形方程式に帰着し、コスト関数の重ね合わせが可能といった有用な性質をもつことから、このような制御問題は線形可解 (linearly solvable) マルコフ決定過程と呼ばれる [49]。なお、(10.21) は Z の定数倍に依存しない。これは、評価関数 ℓ の全状態に一様な定数を加える自由度に対応する。

(2) 逆強化学習 与えられた制御対象およびコスト関数に対して最適制御則を求めてきたが、制御対象および最適制御則が与えられたときに、対応するコスト関数を求める問題を考えよう。例えば、LQR 問題において、フィードバックゲイン行列が与えられたときに、そのゲインが最適となるコスト関数の係数行列を求める問題である。強化学習と同様に、モデルを用いることなくこの問題を解くことを逆強化学習 (inverse reinforcement learning) と呼ぶ。実用的にも、熟練者の技能が観測できる場合に、それを解釈したり引き継ぐ際に有用であろう。素朴に考えると、コスト関数は与えられた制御則により小さくなる関数であると考えられる。例えば、問題 6.1.1 に対する最適制御則 $u = \pi(x)$ 、状態価値関数 $V(x)$ 、遷移確率密度関数が既知の場合にステージコスト $\ell(x, u)$ は求められるであろうか。定理 6.1.4 により $\arg \min_u (\ell + \beta \mathcal{A}_x[V])(x, u) = \pi(x)$ となる関数であるが、これだけでは ℓ は一意に定まらないことがこの問題の本質的な困難さである（図 10.1）。この問題に対して、最大エントロピー原理にもとづいたり、与えられた制御問題の解は比較的簡単に求められるという前提でコスト関数に関する勾配法を用いる手法などが知られている。

ところで、最適制御入力値 $\pi(x)$ は $\ell(x, \cdot)$ に関して一点のみ ($u = \pi(x)$) における値の情報しか与えないことが、図 10.1 における ℓ の非決定性の要因である。それでは、最適制御則が確率的状態フィードバックで与えられる場合はどうなるであろうか。この問い合わせるために、問題 10.1.4 に対する逆強化学習を考える。状態コスト ℓ が未知であり、制御をしない場合の状態 j から i への

図 10.1: 逆強化学習（同一の $\pi(x)$ となる評価関数）

遷移確率 $\Psi^0(i|j) := (P^0)_{ij}$ は既知、最適制御則 (10.21) のもとでの軌道データ $t := (x_0^\pi, x_1^\pi, \dots, x_{\bar{k}}^\pi)$ が利用可能であるとする。ここで目標は最適軌道データ t から状態コスト関数 ℓ 、適合関数 Z 、価値関数 V 、最適制御則 π を推定することである。それぞれの関係式 (10.21), (10.22), (10.23) があるため、どれか一つを推定するだけで十分であるが、これから説明するように V の推定は凸最適化として定式化できる [50]。

まず、最適制御則 π が (10.21) のように適合関数 Z でパラメetrizeされていると見なす。最適軌道 $x_{0:\bar{k}}^\pi$ の負の対数尤度は

$$\begin{aligned} L[Z] &:= -\log \varphi_{x_{0:\bar{k}}^\pi}(t) \\ &= -\log \left(\varphi_{x_0^\pi}(x_0^\pi) \prod_{k=1}^{\bar{k}} \Psi^\pi(x_k^\pi | x_{k-1}^\pi) \right) \\ &= -\sum_{k=1}^{\bar{k}} \log(Z)_{x_k^\pi} + \sum_{k=1}^{\bar{k}} \log \underbrace{\sum_{i=1}^n \Psi^0(i|x_{k-1}^\pi)}_{\text{分母の } Z^\top(P^0)_{:,j} \text{ から}} (Z)_i + \text{定数} \end{aligned}$$

で与えられる。ここで、 $L[Z]$ は Z の要素に関して凸になっていないが、

$$a(i) := \#\{k \in [\bar{k}] : x_k^\pi = i\}, \quad b(j) := \#\{k \in [\bar{k}] : x_k^\pi = j\} \quad (10.29)$$

を数えて、 V を変数として書き換えると

$$L[V] = \beta \sum_{i=1}^n a(i)(V)_i + \sum_{j=1}^n b(j) \log \sum_{i=1}^n (P^0)_{ij} \exp(-\beta(V)_i)$$

は V の要素について凸になり、効率良く最尤推定が計算できる。

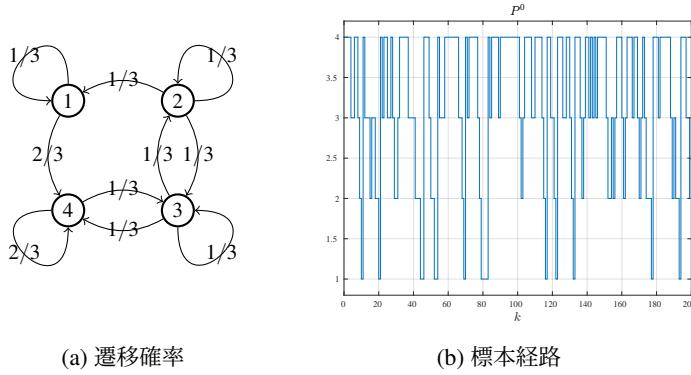


図 10.2: マルコフ連鎖

例 10.1.6 (KL 制御). 図 10.2a の遷移確率をもつマルコフ連鎖に対応する標本経路を図 10.2b に示す ($p_k \rightarrow [0.076 \quad 0.153 \quad 0.307 \quad 0.461]^\top$). これに対して $\ell = \sigma [0 \quad 1 \quad 2 \quad 3]$ のステージコストをもつ KL 制御問題の最適遷移確率図 10.3a を考える. σ が大きくなるにつれて相対的に遷移確率を修正するコストが低くなるため, 小さい番号の状態に滞在する時間が長くなる制御則が得られている ($p_k^\pi \rightarrow [0.595 \quad 0.157 \quad 0.132 \quad 0.114]^\top$ ($\sigma = 1$) および $[0.768 \quad 0.092 \quad 0.074 \quad 0.065]^\top$ ($\sigma = 1.2$)). また, 図 10.4a に示す図 10.3a の標本経路のデータと図 10.2a の遷移確率を用いて, 逆強化学習をおこなった結果を図 10.4b に示す ($\ell_1 = 0$ に固定した).

△

演習 10.2. 例 10.1.6 の結果を計算機を用いて確認せよ.

△

10.2 漸近拳動

10.2.1 ペロン・フロベニウスの定理

本節では, ペロン・フロベニウスの定理を用いてマルコフ連鎖の確率分布の漸近拳動に関する重要な性質を紹介し, マルコフ連鎖の収束性を示す (定理 10.2.5).

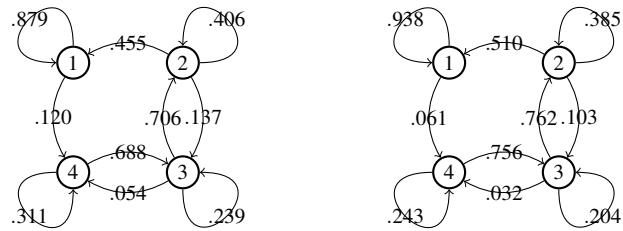
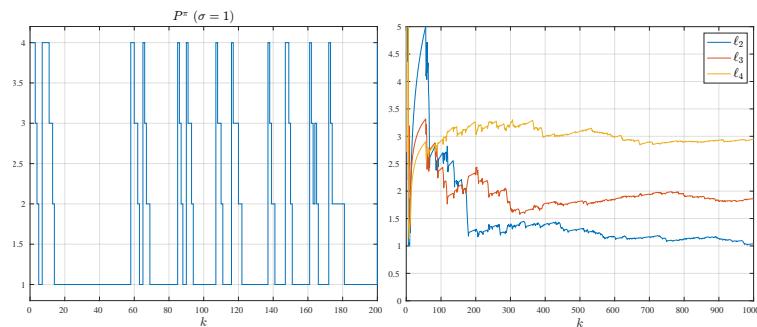
(a) $\sigma = 1$ (b) $\sigma = 1.2$

図 10.3: 最適遷移確率



(a) 標本経路

(b) 推定ステージコスト

図 10.4: 逆強化学習

定義 10.2.1. 非負行列 $A \in \mathbb{R}^{n \times n}$ が既約(irreducible)であるとは、任意の $i, j \in \llbracket n \rrbracket$ に対して、ある $k_1(i, j) \in \mathbb{N}$ が存在し、 $(A^{(k_1(i, j))})_{ij} > 0$ が成り立つことをいう。

△

定義 10.2.2. 遷移確率行列 $P \in \mathbb{R}^{n \times n}$ をもつマルコフ連鎖の状態 $i \in \llbracket n \rrbracket$ について、ある $k_2(i) \in \mathbb{N}$ が存在して、任意の $k \geq k_2$ に対して $(P^k)_{ii} > 0$ となるとき、状態 i は非周期的(aperiodic)であるという。△

遷移確率行列が既約であるならば、任意の状態 j から i への遷移が有限時間で起こる確率が正であることを要求しており、図 10.2a のようなグラフ表現では、任意のノード間に方向付き経路が存在することに対応する。非周期性は、「 $x_0 = i$ のとき、 k が 3 の倍数のときのみ $\mathbb{P}(x_k = i) > 0$ である」といった状況を排除している。

証明は割愛するが、マルコフ連鎖の解析においては、つきの定理が中心的な役割を果たす。

定理 10.2.3 (ペロン・フロベニウスの定理). 非負正方形 A に対して、

1. A は非負の実固有値をもち、そのうち最大のものを $\bar{\lambda}$ (A のフロベニウス根と呼ぶ) とおけば、 $\bar{\lambda}$ に対応する非負固有ベクトルが存在し、
2. A の任意の固有値 λ は $|\lambda| \leq \bar{\lambda}$ を満たす。

また、 A が既約な非負行列ならば、 $\bar{\lambda}$ は単純固有値（重複度が 1）である。さらに A が正行列ならば、 $\bar{\lambda}$ を除いた A の全ての固有値 λ は $|\lambda| < \bar{\lambda}$ を満たす。△

命題 10.2.4. 確率行列 $P \in \mathbb{R}^{n \times n}$ のフロベニウス根は 1 である。△

証明. P の任意の固有値 λ と対応する固有ベクトル v について、 $Pv = \lambda v$ であるから、

$$|\lambda \cdot (v)_i| = \left| \sum_{j=1}^n (P)_{ij} (v)_j \right| \leq \sum_{j=1}^n (P)_{ij} |(v)_j|, \forall i \in \{1, \dots, n\}$$

が成り立つ。特に、 $|(v)_i|$ を最大にする $i = i^*$ について

$$\sum_{j=1}^n (P)_{i^*j} |(v)_j| \leq \sum_{j=1}^n (P)_{i^*j} |(v)_{i^*}| = |(v)_{i^*}|$$

であり, $|(v)_{i^*}| \neq 0$ であるから, $|\lambda| \leq 1$ が成り立つ. ところで, $P^\top \mathbb{1} = \mathbb{1}$ より P は固有値 1 をもつから, $\bar{\lambda} = 1$ は P のフロベニウス根である. \square

遷移確率行列 $P \in \mathbb{R}^{n \times n}$ をもつマルコフ連鎖において,

$$Pp = p \quad (10.30)$$

を満たす確率ベクトル p をマルコフ連鎖の定常分布 (stationary distribution) もしくは不变分布 (invariant distribution) と呼ぶ. 命題 10.2.4 とペロン・フロベニウスの定理より, 固有値 1 に対応する P の非負固有ベクトル $v^* \in \mathbb{R}^n$ をとることができ, 規格化により確率ベクトル

$$p_{\text{st}} := \frac{v^*}{\sum_{i=1}^n (v^*)_i} \quad (10.31)$$

を構成することができる. すると $Pp_{\text{st}} = p_{\text{st}}$ を満たすから, p_{st} は定常分布である. P が既約なとき, 固有値 1 は単純であるから, p_{st} は v^* のとり方によらず一意に定まる. さらに, 収束性に関してつぎの定理が成り立つ (証明は B.4 を参照).

定理 10.2.5 (マルコフ連鎖の収束). 既約な遷移確率行列 $P \in \mathbb{R}^{n \times n}$ をもつマルコフ連鎖のある状態 $i \in [n]$ が非周期的であるとする. このとき, 任意の初期分布 p_0 に対して,

$$\lim_{k \rightarrow \infty} P^k p_0 = p_{\text{st}} \quad (10.32)$$

が成り立つ. \triangleleft

10.2.2 定常状態と平衡状態

入力をもたない確定システム $x(k+1) = f(x(k))$ に対して $f(x_{\text{eq}}) = 0$ を満たす x_{eq} は平衡点と呼ばれ, $x(k) = x_{\text{eq}}$ ならば $x(l) = x_{\text{eq}}, \forall l \geq k$ が成り立つ. 本節では, 確率システムに対する同様の性質を考察する.

定義 10.2.6 (定常確率分布). 時刻に依存しない遷移確率度関数 Ψ をもつ \mathbb{X} -値マルコフ過程 x_i に対して, \mathbb{X} 上の確率密度関数 φ_{st} が

$$\varphi_{st}(x') = \int_{\mathbb{X}} \Psi(x'|x) \varphi_{st}(x) dx \quad (10.33)$$

を満たすとき, φ_{st} を x_k の**定常確率分布** (stationary probability distribution) と呼ぶ. \triangleleft

定理 3.1.7 を踏まえると自然な定義であり, (10.33) は (10.30) に対応する条件である. この条件が成り立つとき, $\varphi_{x_k} = \varphi_{st}$ ならば, $\varphi_{x_l} = \varphi_{st}, \forall l \geq k$ である. ここですべての標本経路が一定値をとりつづける $x_l = x_k$ という意味ではなく, あくまで確率密度関数の不变性であることに注意する. 平衡点が関数 f の不動点であるのに対して, 定常確率分布は (10.33) の解であるため, 与えられたマルコフ過程の定常確率分布を具体的に求めることは (存在したとしても) 一般的に困難である. 非周期性などの諸概念に加え緊密性などが重要な役割を果たすが, 詳細は [51]などを参照されたい.

本章の残りでは, 定常確率分布のうちつぎの性質をもつものに注目した議論をおこなう.

定義 10.2.7 (詳細釣り合い条件). 時不变 \mathbb{X} -値マルコフ過程 x_i に対して, ある \mathbb{X} 上の確率分布 φ_{eq} が存在し,

$$\Psi(x'|x) \varphi_{eq}(x) = \Psi(x|x') \varphi_{eq}(x'), x \neq x' \quad (10.34)$$

を満たすとき, x_i は詳細釣り合い条件を満たすといい, φ_{eq} を**平衡分布** (equilibrium distribution) と呼ぶ. \triangleleft

(10.34) の両辺を x で積分すると (10.33) が得られるため, 平衡分布は定常分布である. (10.34) は, 状態が平衡状態にあるとき, x から x' への移動量はその逆方向の移動量に等しいことを要求しており, 見方を変えると時間的に可逆であることを意味している.

一つ応用例を見てみよう。ベイズ統計の活用においては、事後分布 $\varphi_{x|y=y}(x) = \varphi_{(x,y)}(x, y)/\varphi_y(y)$ の性質を調べることは中心的な課題である。ここで、 $\varphi_{(x,y)}$ が与えられている場合、分子は代入操作だけであるが、分母の値を求める周辺化に必要な積分操作は困難である場合が多い。したがって、与えられた確率密度関数 φ_d の規格化定数が未知の場合に、これを定常分布として持つような時不变のマルコフ過程を与えることは、応用上も非常に有用である。

定理 10.2.8 (メトロポリス・ヘイスティングス法). 非負値関数 $\psi_d : \mathbb{X} \rightarrow \mathbb{R}_+$ および $q : \mathbb{X} \times \mathbb{X} \rightarrow \mathbb{R}_+$ が

$$c_d := \int_{\mathbb{X}} \psi_d(x) dx < \infty, \quad (10.35)$$

$$q(x|x') = q(x'|x), \quad \int_{\mathbb{X}} q(x'|x) dx' = 1 \quad (10.36)$$

を満たすとする。このとき、マルコフ過程 x_k を

1. x_k^+ は $\varphi_{x_k^+}(x'|x_k=x) = q(x'|x)$, $r_k \sim \text{Uni}([0, 1])$ の独立な確率変数,
2. $r_k \leq \alpha_k := \min\{1, \psi_d(x_k^+)/\psi_d(x_k)\}$ ならば $x_{k+1} = x_k^+$, それ以外の場合は

$$x_{k+1} = x_k$$

により定義する。このとき、 $\varphi_d(x) := \psi_d(x)/c_d$ は x_k の平衡分布である。 \square

証明. 任意の $x \neq x'$ に対して,

$$\varphi_{x_{k+1}}(x'|x_k=x) = q(x'|x) \min \left\{ 1, \frac{\varphi_d(x')}{\varphi_d(x)} \right\} \quad (10.37)$$

$$\varphi_{x_{k+1}}(x'|x_k=x) \varphi_d(x) = q(x'|x) \min \{ \varphi_d(x), \varphi_d(x') \} \quad (10.38)$$

である。この右辺は x, x' に関して対称であることに注意すると $\varphi_{x_{k+1}}(x'|x_k=x) \varphi_d(x) = \varphi_{x_{k+1}}(x|x_k=x') \varphi_d(x')$ が成り立つ。 \square

最後に、所望の平衡分布をもつマルコフ連鎖の構築を工学利用する例として、つぎの興味深いアルゴリズムを紹介して本章を締めくくる。

例 10.2.9 (センサーネットワーク配置問題). 図 10.5a の領域 $P := \llbracket 100 \rrbracket \times \llbracket 100 \rrbracket$ があり, 各地点 $z \in P$ において重要度 $V(z)$ が与えられている. ここに $j \in \llbracket 12 \rrbracket$ でラベル付けられた 12 台のセンサーを配置する. ドローンによる環境のモニタリングや災害時の遭難者の探索などを想定すればよい. ただし, 各センサーは, 自身を中心とした半径 10 の円内しか観測できないとする. すなわち, センサー j の位置を $x^j \in P$ とおくと, 配置 $x = (x^1, \dots, x^{12})$ に対して $P(x) := \{z \in P : \exists j \text{ s.t. } \|z - x^j\| \leq 10\}$ がセンサー群全体で観測可能な領域である. また, この配置の評価値

$$\ell(x) := \sum_{z \in P(x)} V(z) \quad (10.39)$$

をできるだけ大きくする配置 $x \in \mathbb{X} := P^{12}$ を求めたい. 組み合わせ総数は 10^{48} 通りなので, 網羅的な評価による最適化は非現実的である.

そこで, 与えられた配置 x に対して, 位置を更新するセンサー $i \in \llbracket 12 \rrbracket$ を等確率で一つだけ選択し, さらに移動方向も上下左右のいずれかから等確率で選択することで, 遷移先候補配置 $\hat{x}^i \in \{x^{i\uparrow}, x^{i\downarrow}, x^{i\leftarrow}, x^{i\rightarrow}\} \subset \mathbb{X}$ を定める[†]. ただし, 評価値の低い配置に移動する確率を下げるため, 確率 $e^{\beta \ell(\hat{x}^i)} / (e^{\beta \ell(x)} + e^{\beta \ell(\hat{x}^i)})$ で現在の位置 x にとどまるものとする. つまり, 時刻 k での配置を x_k とおくと,

$$\mathbb{P}(x_{k+1} = x^{i\leftarrow} | x_k = x) = \frac{1}{12} \times \frac{1}{4} \times \frac{e^{\beta \ell(x^{i\leftarrow})}}{e^{\beta \ell(x^{i\leftarrow})} + e^{\beta \ell(x)}} \quad (10.40)$$

などで記述されるマルコフ過程に対応する. なお, 位置を更新するセンサーは一時刻につき一台のみであることに注意する. 図 10.5b にすべてのセンサーが $(10, 10) \in P$ に初期配置されている状態からの標本経路を示すが, センサーが重要度の高い領域を被覆するように移動していることがわかる. □

あきらかに任意の二配置間の遷移がおこる確率は正である. さらに,

$$\mathbb{P}_\infty(x) \propto e^{\beta \ell(x)} \quad (10.41)$$

[†] 配置 x に対して, $x^{i\leftarrow}$ はセンサー i のみを左に一マス移動させた際に得られる配置をあらわす. また, x が領域の左端の場合は $x^{i\leftarrow} = x$ などとする.

で与えられる \mathbb{X} 上の確率分布が平衡分布であることが簡単な計算で確認できる。したがって、この分布は定常分布であり、このマルコフ過程は \mathbb{P}_∞ に収束する。また、(10.41) は $\ell(x)$ が大きい配置 x ほど高確率で実現するという意味で、当初の目的を満足する。

ここで、各時刻の更新には (10.40) の計算が必要であるため、すべてのセンサー配置 x の情報を把握する集中管理者が必要であるように見える。しかし、位置を更新するセンサー i にはその旨を伝える必要があるが、(10.40) の計算に必要なのは差 $\ell(x^{i \leftarrow}) - \ell(x)$ の値だけであり、観測円内の情報だけから算出できる（遠く離れた場所が被覆されているかどうかは差の値には影響しない）。つまり、例 10.2.9 のアルゴリズムは分散的に実装することができる。さらに踏み込んでポテンシャルゲームの枠組みでナッシュ均衡との関係も議論することができるので、興味がある場合は文献 [52]などを参照してほしい。なお、(10.40) の更新則は評価値が減少する遷移も許容する。これに対して、評価値が減少する遷移は許容しない場合の挙動を図 10.5c に示す。この場合、 $k = 4226$ 以降はこれ以上移動はおこらない局所最適配置に陥っている。

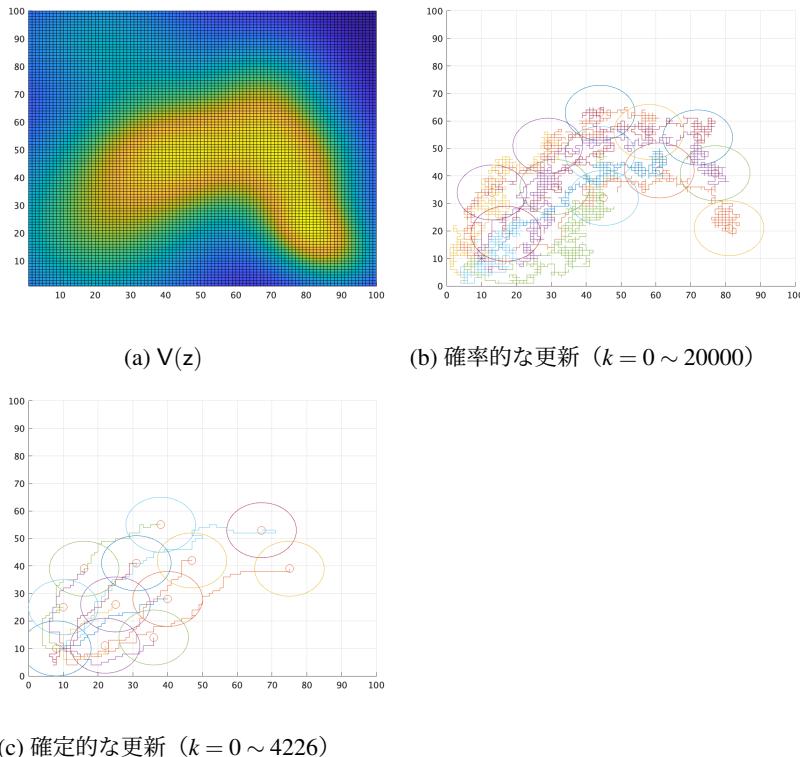


図 10.5: ポテンシャルゲームによるセンサーネットワーク配置（各センサーの軌道と最終時刻における位置（小円）および観測可能領域（大円）).

11|| 数理最適化 (16)

制御や推定問題がベルマン方程式や誤差最小化問題に帰着することを見てきたが、これらを実際に解くには、なんらかの数理最適化が不可欠である。一方で、システム制御分野と統計的学習分野では、表 11.1 のように最適化手法に対する考え方方が根本的に異なる。本章では、11.1 節でシステム制御分野、11.2 節にて統計的学習分野の代表的手法をそれぞれ紹介するので、分野間の発想の違いを実感してほしい。

11.1 半正定値計画問題

システム制御理論においては、大域的最適解をある程度安定して求められる最適化手法が好まれる傾向がある。その結果として、解きたい問題を凸最適化問題に帰着させることが、目標として設定されることが多い。とりわけ、線形システムの制御問題は半正定値計画問題（Semi-positive Definite Programming; SDP）もしくは線形行列不等式（Linear Matrix Inequality; LMI）との相性がよく、広い観点から研究がおこなわれてきた [53,54]。そこで本節では、半正定値計画問題にもとづく制御系設計の典型的な手法を紹介する。

11.1.1 ベルマン方程式の不等式評価

ふたたび、本書で主たる対象とする状態集合 \mathbb{X} が \mathbb{R}^{n_x} もしくはその部分集合となる場合を考える。この場合、状態集合が無限集合となることから、LQR 問題などごく限られた問題を除いて、ベルマン方程式や関連する不等式が容易に

表 11.1: 数理最適化

	システム制御	統計的学習
最適性	大域的最適解	局所最適解, 確率的最適性
最適化問題	凸最適化問題	勾配法, 標本近似
問題例	二次計画問題, 半正定値計画問題	確率勾配法
計算時間	多項式オーダー	線形オーダー

解けることはない。したがって、現実的な計算量で解を得るには何らかの工夫が必要となるが、11.1.1節では問題4.1.4の最適制御問題に対して、ダイナミクスや評価関数が多項式で表現できる非線形性を有する場合に限定した議論をおこなう（区分的多項式でも同様の議論が可能である）。

仮定 11.1.1. $m_r(x, u)$ を x, u の要素の次数 r 以下の単項式からなるベクトルとし、さらに u を含まない項を抜き出したベクトルを $m_r(x)$ と表記する。このとき、偶数 $r, r' < r$ と整数 $\underline{r} \leq r$ に対して、

$$\ell(x, u) = c_\ell^\top m_r(x, u), \quad \ell_f(x) = c_f^\top m_r(x) \quad (11.1)$$

$$\mathcal{C} := \{(x, u) : x \in \mathbb{X}, u \in U(x)\} = \{u : c_h^\top m_{r'}(x, u) < 0\} \quad (11.2)$$

$$\mathcal{A}_x[m_r] := \int_{\mathbb{X}} m_r(x') \Psi(x'|x, u) dx' = c_\psi m_r(x, u) \quad (11.3)$$

$$\mathbb{E}_{x_0}[m_r(x_0)] = c_0 \quad (11.4)$$

が成り立ち、 $c_\ell, c_f, c_h, c_\psi, c_0$ は既知である。 \triangleleft

たとえば、 $x := [x_1, x_2]^\top$ が 2 次元、 u がスカラーであるとすると、

$$m_2(x, u) = [1, x_1, x_1^2, x_2, x_1 x_2, x_2^2, u, u x_1, u x_2, u^2]^\top, \quad (11.5a)$$

$$m_2(x) = [1, x_1, x_1^2, x_2, x_1 x_2, x_2^2]^\top \quad (11.5b)$$

である。この記法のもと、任意の r 次以下の多項式は、実数ベクトル c と $m_r(x, u)$ の内積 $c^\top m_r(x, u)$ と表現できる。 (11.2) は制御入力の状態制約が多項式の不等式で与えられることを仮定している。また、 $x_{k+1} = f(x_k, u_k, v_k)$ の右辺が x_k, u_k に関して多項式ならば、任意の多項式ベクトル $m(x)$ に対して $\int m(f(x, u, v)) \varphi_v(v) dv$ も多項式であるように、 (11.3) はダイナミクスの多項式性を表現している。

(1) 多項式計画問題 ここでは、(10.4)に相当する最適化問題を導出する。得られる結果は必ずしも実用的に優れているわけではないが、連続値の関数不等式の取り扱い方の好例である。系 4.2.2 より、

$$\check{V}(\bar{k}, \mathbf{x}) \leq \ell_f(\mathbf{x}), \forall \mathbf{x} \in \mathbb{X}, \quad (11.6a)$$

$$\check{V}(k, \mathbf{x}) \leq (\ell + \mathcal{A}_x[\check{V}])(k, \mathbf{x}, \mathbf{u}), \forall k \in \llbracket \bar{k} \rrbracket, \mathbf{x} \in \mathbb{X}, \mathbf{u} \in U_k(\mathbf{x}), \quad (11.6b)$$

を満たす任意の \check{V} に対して、 $V^* \geq \check{V}$ が成り立つ。したがって、適切な関数の集合 \mathcal{F} に対して、

$$\text{maximize}_{\check{V} \in \mathcal{F}} \mathbb{E}_{x_0} [\check{V}(0, x_0)] \quad \text{subj. to} \quad (11.6) \quad (11.7)$$

を解けば、最適制御性能の下限が求まる。ここで、(11.7)における \mathcal{F} として r 次以下の多項式からなる関数を考えるとつぎの結果が得られる。

定理 11.1.2. 仮定 11.1.1 のもとでの問題 4.1.4 を考える。このとき、

$$-\eta_{\bar{k}}^\top m_r(\mathbf{x}) + c_f^\top m_r(\mathbf{x}) = m_{r/2}(\mathbf{x})^\top M_{\bar{k}} m_{r/2}(\mathbf{x}), \quad (11.8a)$$

$$\begin{aligned} -\eta_k^\top m_r(\mathbf{x}) + (c_\ell^\top + \eta_{k+1}^\top c_\psi) m_r(\mathbf{x}, \mathbf{u}) + c_h^\top m_{r'}(\mathbf{x}, \mathbf{u}) \\ \times m_{(r-r')/2}(\mathbf{x}, \mathbf{u}) \tilde{M}_k m_{(r-r')/2}(\mathbf{x}, \mathbf{u}) = m_{r/2}(\mathbf{x}, \mathbf{u})^\top M_k m_{r/2}(\mathbf{x}, \mathbf{u}), \end{aligned} \quad (11.8b)$$

$$M_k \succeq O, \tilde{M}_k \succeq O \quad (11.8c)$$

を満たす η_k, M_k, \tilde{M}_k が存在するならば、任意の $u_:$ に対して $J(u_:) \geq \eta_0^\top c_0$ である。 ◇

証明. $\check{V}(k, \mathbf{x}) = \eta_k^\top m_r(\mathbf{x})$ とおくと、 $\mathbb{E}[\check{V}(0, x_0)] = \eta_0^\top c_0$ であり、

$$\mathcal{A}_x[\check{V}](k, \mathbf{x}, \mathbf{u}) = \int_{\mathbb{X}} \check{V}(k+1, \mathbf{x}') \Psi(\mathbf{x}' | \mathbf{x}, \mathbf{u}) d\mathbf{x}' = \eta_{k+1}^\top c_\psi m_r(\mathbf{x}, \mathbf{u})$$

である。よって、(11.8a) から (11.6a)、また (11.8b) から (11.6b) がそれぞれ成り立つ。 □

$M_{\bar{k}}$ の半正定値性より $M_{\bar{k}} = \sum_i \xi_i \xi_i^\top$ と分解できるため、(11.8a) の右辺は多項式 $\xi_i^\top m_{r/2}(x)$ の二乗の和となる。このように多項式の大域的非負性を二乗和に分解できるという十分条件に置き換える手法を二乗和緩和 (Sum-of-Squares relaxation) と呼ぶ [55]。与えられた η_k に対して、(11.8a) および (11.8b) を満たす行列 M_k, \tilde{M}_k は必ず存在するが、一意ではない（たとえば $1 \times x_1^3 = x_1^1 \times x_1^2$ に対応する自由度）。この自由度を活用して、(11.8c) の半正定値性を満たす行列を求める構造となっている。この制約のもとで $\eta_0^\top c_0$ を最大化する問題は半正定値計画問題と呼ばれ効率的に解くことができる問題のクラスである。

無限区間の制御問題の不等式（系 6.1.6）も、同様に考えることができる。有限区間の場合と比較して状態価値関数は時間変数を引数としてもたないため、 $\check{V}(x) := \eta^\top m_r(x)$ とあらわされる関数および

$$\eta^\top m_r(x) \leq c_\ell^\top m_r(x, u) + \beta \eta^\top c_\psi m_r(x, u) \quad (11.9)$$

などに対応する不等式拘束を考え、同様の最大化問題を解けばよい。

(2) 一般化モーメント問題 実数値確率変数 x および（区分的）多項式 h_i, g に対して、 $\mathbb{E}[h_i(x)], i = 1, 2, \dots$ の値が与えられたとき、 φ_x の形状は何かの制約をうける（定理 3.1.13 も参照）が、 $\mathbb{E}[g(x)]$ はどのような値を取りうるだろうか。このような問題を一般化モーメント問題（generalized moment problem）と呼ぶ。例えば、 $b_d := \mathbb{E}[x^d] \geq 0, d = 4, 6$ の値が与えられるときに、 $g(x) = x^5$ の期待値が知りたければ、

$$\mathbb{E} \begin{bmatrix} \begin{bmatrix} x^2 \\ x^3 \end{bmatrix} & [x^2, x^3] \end{bmatrix} = \begin{bmatrix} b_4 & \mathbb{E}[x^5] \\ \mathbb{E}[x^5] & b_6 \end{bmatrix} \succeq O \quad (11.10)$$

から $(\mathbb{E}[x^5])^2 \leq b_4 b_6$ である。以下では、この一般化モーメント問題の枠組みで、線形計画問題 (10.11) に関する議論を再訪する。

素朴に考えると (10.11) に対応する決定変数は確率密度関数 $\{\varphi_{(x_k, u_k)}^\pi\}_k$ となるが、ここでは期待値に注目する。まず、

$$\zeta_k := \mathbb{E}^\pi[m_r(x_k, u_k)] \quad (11.11)$$

を確率変数 (x_k^π, u_k^π) のモーメントベクトル, $\mathbb{E}^\pi[m_{r'/2}(x_k, u_k)m_{r'/2}(x_k, u_k)^\top]$ をモーメント行列と呼ぶ。ここで、モーメント行列の半正定性に注目すると、つぎの結果が得られる。

定理 11.1.3 (モーメント行列の正定性にもとづく評価)。仮定 11.1.1 のもとでの問題 4.1.4 に対して,

$$m_r(x) = \mathcal{M}_1(m_r(x, u)) \quad (11.12)$$

$$m_{r/2}(x)m_{r/2}(x)^\top = \mathcal{M}_2(m_r(x)) \quad (11.13)$$

$$c_h^\top m_{r'}(x, u)m_{(r-r')/2}(x, u)m_{(r-r')/2}(x, u)^\top = \mathcal{M}_3(m_r(x, u)) \quad (11.14)$$

となる線形写像 $\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_3$ をとする[†]。このとき, $J(\pi) < J^*$ となる制御則 π が存在するならば,

$$\mathcal{M}_1(\zeta_{k+1}) = c_\psi \zeta_k, \mathcal{M}_1(\zeta_0) = c_0 \quad (11.15a)$$

$$\mathcal{M}_2(\zeta_k) \succeq O, \quad (11.15b)$$

$$\mathcal{M}_3(\zeta_k) \succeq O, \quad (11.15c)$$

および

$$J(\zeta) := \sum_{k=0}^{\bar{k}-1} c_\ell^\top \zeta_k + c_f^\top c_\psi \zeta_{\bar{k}-1} < J^* \quad (11.16)$$

を満たす $\{\zeta_k\}$ が存在する. \triangleleft

証明. この π に対して, (11.11) の ζ_k が条件を満たすことを示す。まず,

$$\begin{aligned} \mathbb{E}^\pi[m_r(x_{k+1})] &= \mathbb{E}^\pi \left[\int_{\mathbb{X}} m_r(x') \Psi_k(x'|x_k, u_k) dx' \right] \\ &= \mathbb{E}^\pi [c_\psi m_r(x_k, u_k)] = c_\psi \zeta_k \end{aligned} \quad (11.17)$$

[†] 具体的には, $m_r(x)$ の各成分は $m_r(x, u)$ の成分に含まれるので, 例えば, (11.5)においては \mathcal{M}_1 は $[I \ O]$ を左からかける操作とすればよい。同様に $\mathcal{M}_2, \mathcal{M}_3$ も適切な成分を抽出し並べる操作である。

176 11. 数理最適化(16)

から (11.15a) が成り立つ。また, $\mathcal{M}_2(\zeta_k) = \mathbb{E}^\pi[m_{r/2}(x_k)m_{r/2}(x_k)^\top] \succeq O$ であり, (11.15c) も同様に確かめられる。最後に, (11.16) の $J(\zeta)$ は $J(\pi)$ に等しい。□

この定理の対偶をとると, 条件を満たす ζ が存在しないならば $J(\pi) < J^*$ となる π は存在しない。したがって, こうした ζ の存在性が失われる最大の J^* , つまり

$$\underset{\{\zeta_k\}_{k=0}^{L-1}}{\text{minimize}} \quad J(\zeta) \quad \text{subj. to (11.15)} \quad (11.18)$$

の最小値は最適制御性能の下限となる。ここで, 次数 r の増加とともに制約条件が増えるため最小値の値は増加し, よりよい下界が得られる構造となっている。モーメント行列の正定性にもとづく正定値計画問題の代わりに, チェビシェフの不等式を用いて線形計画問題に帰着することもできる。これらの詳細な比較および次数をあげた場合の近似誤差の評価は [56] が詳しい。

注意 11.1.4. すべての次数のモーメントが一致しても分布が一致するとは限らない [57, p. 43]。また, 高次のモーメントを低次のモーメントを用いて近似するモーメントクロージャと呼ばれる手法もある。△

演習 11.1. 無限区間の制御問題に対して,

$$\zeta := \mathbb{E}^\pi \left[\sum_{k=0}^{\infty} \beta^k m_r(x_k, u_k) \right] \quad (11.19)$$

に注目すると, 評価指標は $J(\pi) = J(\zeta) := c_\ell^\top \zeta$, 不等式制約は $\mathcal{M}_i(\zeta) \succeq O$, $i = 2, 3$, (11.15a) に対応する等式制約は

$$\mathcal{M}_1(\zeta) = c_0 + \beta c_\psi \zeta \quad (11.20)$$

により与えられることを示せ。△

11.1.2 状態制約付き線形制御

例 5.3.3 において, 線形確定システムの最適制御問題を二次計画問題に帰着させて解を与えた。本節では,

$$x_{k+1} = Ax_k + Bu_k + v_k, \mathbb{E}[v_k] = 0 \quad (11.21)$$

で与えられる線形確率システムの制御問題を考える。確定的な入力では状態の分散に影響を与えられず（定理 3.2.1），システムの安定化などはできないので（例 4.1.1），ここでは線形状態フィードバック制御則

$$u_k = K_k x_k + z_k, \text{Var}[z_k] = Z_k \quad (11.22)$$

を求めよう。このとき， $\Sigma_k := \mathbb{E}[x_k x_k^\top]$ は，

$$\Sigma_{k+1} = (A_k + B_k K_k)^\top \Sigma_k (A_k + B_k K_k) + B_k Z_k B_k^\top + \text{Var}[v_k] \quad (11.23)$$

にしたがう。ただし， Σ_k, K_k, Z_k を決定変数とする最適化問題は， Σ_k と K_k の積が現れるため，高速に解くことは困難である。そこで，

$$M_k := K_k \Sigma_k K_k^\top + Z_k, P_k := K_k \Sigma_k \quad (11.24)$$

により変数変換する。シューフ補行列の性質より，

$$\begin{bmatrix} \Sigma & P^\top \\ P & M \end{bmatrix} \succeq O \iff \Sigma \succeq O, (I - \Sigma \Sigma^\dagger) P^\top = O, M - P \Sigma^\dagger P^\top \succeq O \quad (11.25)$$

が成り立ち，定義より $(I - \Sigma_k \Sigma_k^\dagger) P_k^\top = O, M_k - P_k \Sigma_k^\dagger P_k^\top \succeq Z_k \succeq O$ である。したがって，

$$\begin{bmatrix} \Sigma_k & P_k^\top \\ P_k & M_k \end{bmatrix} \succeq O \quad (11.26)$$

を課すと，この変換は全单射である。すなわち，コスト関数が Σ_k, M_k, P_k の線形関数で表現できれば，全体として制御問題は半正定値計画問題となり，効率的に解くことが可能である。例えば，二次形式は

$$\begin{bmatrix} x_k \\ u_k \end{bmatrix}^\top \begin{bmatrix} Q_k & S_k \\ S_k^\top & R_k \end{bmatrix} \begin{bmatrix} x_k \\ u_k \end{bmatrix} = \text{Trace} \left(\begin{bmatrix} Q_k & S_k \\ S_k^\top & R_k \end{bmatrix} \begin{bmatrix} \Sigma_k & P_k^\top \\ P_k & M_k \end{bmatrix} \right) \quad (11.27)$$

3D Visualization of System Evolution

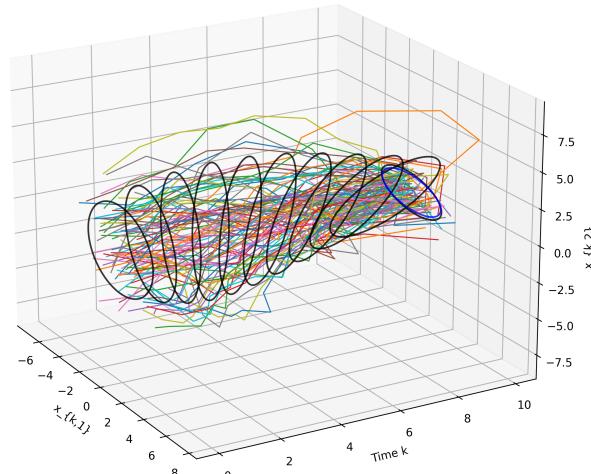


図 11.1: 最適輸送制御

などと表現できる。

他の制約も、適切な不等式で表現できれば、最適化に組み込むことができる。例えば、安全領域から出ないように、行ベクトル C_k に対して $y_k := C_k x_k < 1$ を課したいとする。このとき、確率 1 でこの制約を満たすことは難しい（雑音が非有界であれば不可能）が、代わりに $\mathbb{P}(y_k \geq 1) < \varepsilon$ などと緩和すれば、 y_k の分散 $C_k \Sigma_k C_k^\top$ に関する線形不等式を用いて保証できるだろう。このような事象の生成確率に関する制約を機会制約 (chance constraint) と呼ぶ。

例 11.1.5. 図 11.1



11.2 確率勾配法

前節では、大域的最適解が得られる問題へと緩和されていたものの、11.1.1節の手法は多項式の次数や変数の次元に関して問題のサイズが指数関数的に増大し、11.1.2節の手法は線形システムに限られていた。特に後者は、適切な近似と併用することで、高速かつ信頼性の高い制御入力の決定に役立つ一方で、機械学習分野で扱う変数は画像の全画素値、巨大なニューラルネットワークのパラメータ、様々な環境下で取得されたビッグデータなど、次元が圧倒的に高い。こうした問題に対しては、そもそも大域的最適解を求めるることは現実的ではなく、適切な計算量と精度で局所最適解を得ることが主眼となってくる。典型的には、7章の回帰問題ではなどにおいて、

$$L(\mathbf{p}) := \sum_{s=1}^S L_s(\mathbf{p}), \quad (11.28)$$

$$L_s(\mathbf{p}) := \|f(\mathbf{x}_s; \mathbf{p}) - \mathbf{y}_s\|^2 \quad (11.29)$$

といった形の関数の最適化が頻繁にあらわれていた。本節では、こういった状況で有効な確率勾配法の基礎的な考え方を概説する。

11.2.1 誤差逆伝播

まず、 $f(\mathbf{x}; \mathbf{p}) = \phi(\mathbf{x})^\top \mathbf{p}$ と書けるならば、最小二乗法により大域的最適解が効率的に求められるが、事前にどのような非線形関数 $\phi(\mathbf{x})$ を準備するべきか自明ではなかった。そこで、パラメータ \mathbf{p}_l をもつ単純な非線形関数 $\psi_l(\cdot; \mathbf{p}_l)$, $l = 0, 1, \dots$ を用意して（注意 11.2.2），それらの合成関数を $f(\mathbf{x}; \mathbf{p})$ としたうえで、関数の入れ子構造と \mathbf{p} の調整によって複雑な非線形性の表現を目指す。このとき、 l を層（layer）と呼び、 $\mathbf{p} := \mathbf{p}_{0:\bar{l}}$ を用いて、

$$z_{l+1} := \psi_l(z_l; \mathbf{p}_l), \quad L_s(\mathbf{p}) := z_{\bar{l}+1} \quad (11.30)$$

のようにあらわせる（(11.29) では、 $z_0 = \mathbf{x}_s, z_{\bar{l}} = f(\mathbf{x}_s; \mathbf{p}), \psi_{\bar{l}}(z; \mathbf{p}_{\bar{l}}) = \|z - \mathbf{y}_s\|^2$ ）。

180 11. 数理最適化(16)

もちろん、(11.28) の評価関数が複雑な非凸関数である場合が多いため、大域的最小化解 p を求ることは現実的ではない。そこで、以下では停留点、つまり局所最適解で妥協するが、十分な表現能力（over parametrization）の仮定のもとで非常に大域的最適に近い解が得られるという研究成果もある [58]。停留点を求めるための標準的手法の一つはニュートン法であるが、 p に関する二階偏微分（ヘッセ行列）の計算量が多いため、ほとんど用いられることはない。

そこで、11.2.2 節を通じて勾配法の実装方法を検討するが、11.2.1 節では特に勾配の計算法を議論する。関数 $L_s(p)$ は既知であるため、その p_l に関する勾配を書き下すことは、原理的には可能である。しかし、層数 \bar{l} や p の次元が大きい場合などは、こうした関数を保持するには膨大な記憶容量が必要となり、それに繰り返し値を代入することも効率的ではない。この問題を回避するために、微分の連鎖律

$$\frac{\partial L_s}{\partial p_l}(p) = \underbrace{\frac{\partial \psi_{\bar{l}}}{\partial z_{\bar{l}}}(z_{\bar{l}}(p), p_{\bar{l}}) \cdots \frac{\partial \psi_{l+1}}{\partial z_{l+1}}(z_{l+1}(p), p_{l+1})}_{=: \Gamma_l(p)} \frac{\partial \psi_l}{\partial p_l}(z_l(p), p_l) \quad (11.31)$$

を用いる。ただし、 z_l の p への依存性も陽に表記した。なお、

$$\Gamma_{\bar{l}} := I, \quad \Gamma_{l-1} := \Gamma_l \frac{\partial \psi_l}{\partial z_l}(z_l(p), p_l) \quad (11.32)$$

と再帰的に表現できる。したがって、与えられた p_*, l_* に対して「 $p = p_*$ における L_s の p_{l_*} に対する勾配」を計算するには、(11.30) の逐次代入により $z_l(p_*)$ を $l = 1$ から順に \bar{l} まで計算（順伝搬と呼ばれる）したのち、逆に順伝搬で求めた $z_l(p^*)$ と p^* を用いて、(11.32) の逐次代入により Γ_l を $l = \bar{l}$ から $l_* - 1$ まで逆順に計算（逆伝搬と呼ばれる）し、

$$\frac{\partial L_s}{\partial p_{l_*}}(p_*) = \Gamma_{l_*-1} \quad (11.33)$$

とすればよい。また、この逆伝搬を用いた勾配の計算式(11.33)を用いて評価関数が小さくなるように p を更新する勾配法の実装方法を総じて誤差逆伝播（back propagation）と呼ぶ。

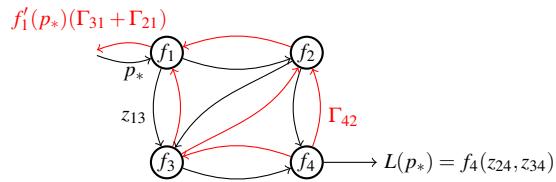


図 11.2: 計算グラフ

行列の積が現れるため、層数 \bar{l} が増える（深層）につれて勾配が 0 になったり（勾配消失）、発散する（勾配爆発）ことがしばしば生じるが、近年はこうした問題を回避する様々な手法が研究されている。また、**計算グラフ** (computation graph) と呼ばれる図 11.2 のような表現を用いて、より複雑な構造をもつ合成関数も系統的に扱うことができる。さらに、繰り返しあらわれるベクトルとヤコビアンの積 (vector-Jacobian product) も、計算の効率化に有用な性質を有する。こうした計算機上で表現された関数の勾配を効率的かつ正確に計算する技術を総じて**自動微分** (Automatic Differentiation あるいは Algorithmic Differentiation; AD) と呼ぶ。

演習 11.2. 図 11.2 は $L(\mathbf{p}) = f_4(f_3(f_1(\mathbf{p}), f_2(f_1(\mathbf{p}))), f_2(f_1(\mathbf{p})))$ の $\mathbf{p} = \mathbf{p}_*$ における勾配を計算するための計算グラフであり、黒が $z_{34} := f_3(z_{13}, z_{23})$ などの順伝搬、赤が $\Gamma_{42} := (\partial f_4 / \partial z_{24})(z_{24}, z_{34})\Gamma_{43}$ などの逆伝搬を表現している。残りの変数も同様に説明せよ。 \triangleleft

注意 11.2.1 (随伴法 (Adjoint method)). 誤差逆伝播が 4.2.3 節で紹介した随伴変数を用いた方法と対応することは明らかである。実際、層の番号 l を時間変数 k と対応させ、(11.30) を状態 z_l 、入力 \mathbf{p}_l 、遷移関数 ψ_l をもつ動的システム、 $z_{\bar{l}}$ を終端状態、 $\Psi_{\bar{l}}$ を終端コスト関数、ステージコストはない、とみなせば、 Γ_l と ζ_k^\top が対応する。 \triangleleft

注意 11.2.2 (ニューラルネットワーク). ニューラルネットワーク (neural networks) の一種である**多層パーセプトロン** (Multi-Layer Perceptron) において

は活性化関数 (activation function) と呼ばれる実数値関数 a を用いて

$$\psi(z, p) := \bar{a}(Wz + b) := \begin{bmatrix} a((Wz + b)_1) \\ \vdots \\ a((Wz + b)_n) \end{bmatrix}, p = (W, b) \quad (11.34)$$

という形で与えられる。活性化関数の選び方に確立された基準はないが、シグモイド関数

$$\varsigma_\alpha(z) := \frac{1}{1 + e^{-\alpha z}} = \frac{\tanh(\alpha z/2) + 1}{2}, \alpha > 0 \quad (11.35)$$

や正規化線形関数 (Rectified Linear Unit; ReLU)

$$\text{ReLU}(z) := \max(0, z) \quad (11.36)$$

は広く採用される。近年は特に、後者を用いた勾配消失や勾配爆発の抑制が注目されている。 ◇

11.2.2 確率近似法

引き続き、(11.28) の停留点、すなわち $\nabla L(p_*) := (\partial L / \partial p)^\top(p_*) = 0$ を満たす p_* を求めるための勾配法を実装すること念頭に議論をすすめる。データ数が大きい場合の課題を検討することから始める。

注意 11.2.3 (ミニバッチ最適化)。前節の結果を用いると、与えられた p_k に対して、 $\nabla L(p_k)$ を計算することは可能である。しかしながら、データ数 \bar{s} が非常に大きい値をとる場合、毎ステップ $\nabla L(p_k)$ を計算することは効率が悪い。そこで確率的に $\lceil \bar{s} \rceil$ の部分集合 S_k を選択し (S_k をミニバッチ、 S_k の要素数をバッチサイズと呼ぶ)、 $\nabla L(p_k)$ の近似値として $\sum_{s \in S_k} \nabla L_s(p_k)$ がしばしば用いられる。また、このミニバッチの確率性より、勾配法が局所最適解に陥ることを抑制する効果もある。 ◇

このように不確実な（確率雑音が付加された）勾配の情報にもとづいて実行される勾配法を一般に 確率勾配法 (Stochastic Gradient Decent; SGD) と呼ぶ。

素朴に考えると、学習率とも呼ばれるステップ幅 $a_k > 0$ を用いて、

$$p_{k+1} = p_k - a_k y_k, \quad y_k := \nabla L(p_k) + e_k \quad (11.37)$$

という方策でうまくいきそうに思える。しかしながら、特に $\nabla L(p_k)$ が十分 0 に近づいたあとなど、観測誤差 e_k の影響で y_k と $\nabla L(p_k)$ の符号が一致しない可能性が残り続けても、適切な意味での収束は望めるのだろうか。以下では、この疑問に関する解析をおこなう。なお、 ∇L に限らない未知方程式 $g(p) = 0$ を解く逐次解法を与えており、この視点からは確率近似法 (stochastic approximation) とも呼ばれる。準備として、つぎの補題を示そう。

補題 11.2.4. 非負の数列 $\{a_k\}_k$, $\{b_k\}_k$, $\{c_k\}_k$ が

$$\sum_k a_k = \infty, \quad \sum_k c_k < \infty \quad (11.38)$$

$$b_{k+1} \leq b_k(1 - a_k) + c_k, \quad \forall k \quad (11.39)$$

を満たすとき、 $\lim_{k \rightarrow \infty} b_k = 0$ が成り立つ。 \triangleleft

証明. (11.39)において厳密な不等式が成り立つ場合は、 c_k を減らすことで、(11.38)を維持したまま等号を成立させることができるために、以下では一般性を失わずに $b_{k+1} = b_k(1 - a_k) + c_k$ を仮定する。このとき、

$$-b_0 \leq b_n - b_0 = \sum_{k=0}^{n-1} (b_{k+1} - b_k) = \sum_{k=0}^{n-1} c_k - \sum_{k=0}^{n-1} a_k b_k \quad (11.40)$$

が成り立つ。ここで、 $\sum c_k < \infty$ より、 $\sum a_k b_k < \infty$ である。ゆえに、右辺の両項ともに収束するため b_k も収束する。また、 $\sum a_k = \infty$ かつ $\sum a_k b_k < \infty$ が成り立つには、この収束値は 0 でなければならない。 \square

この補題を用いて、(11.37)の収束性のための十分条件をつぎのように与えることができる。

定理 11.2.5 (確率勾配法). 時刻独立な確率過程 e_k 、非負の数列 $\{a_k\}$ に対して

$$p_{k+1} = p_k - a_k y_k, \quad y_k := g(p_k) + e_k \quad \text{を考える。このとき,}$$

$$\sum_k a_k = \infty, \sum_k a_k^2 < \infty \quad (11.41)$$

$$g(p_*) = 0, g(p)^\top (p - p_*) \geq \kappa \|p - p_*\|^2, \forall p \quad (11.42)$$

$$\sup_k \mathbb{E} [\|y_k\|^2] < \infty, \quad (11.43)$$

$$\mathbb{E}[e_k] = 0, \sup_k \mathbb{E} [\|e_k\|^2] < \infty \quad (11.44)$$

を満たす定数 κ および $p_* \in \mathbb{R}^n$ が存在するならば, $\lim_{k \rightarrow \infty} \mathbb{E} [\|p_k - p_*\|^2] = 0$ が成り立つ. \triangleleft

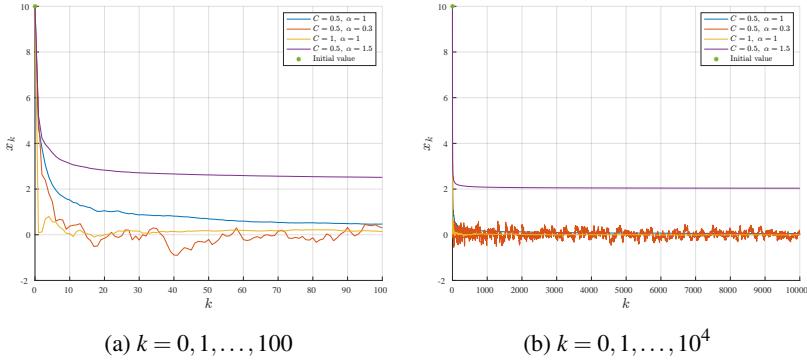
証明. 数列 $b_k := \mathbb{E} [\|p_k - p_*\|^2]$, $\varepsilon_k := \mathbb{E} [\|y_k\|^2]$, $d_k := \mathbb{E} [g(p_k)^\top (p_k - p_*)]$ を定義する. このとき,

$$\begin{aligned} b_{k+1} &= \mathbb{E} [\|p_{k+1} - p_k + p_k - p_*\|^2] \\ &= \mathbb{E} [\|p_{k+1} - p_k\|^2] + 2\mathbb{E} [(p_{k+1} - p_k)^\top (p_k - p_*)] + \mathbb{E} [\|p_k - p_*\|^2] \\ &= \mathbb{E} [a_k^2 \|y_k\|^2] - 2\mathbb{E} [a_k (g(p_k) + e_k)^\top (p_k - p_*)] + b_k \\ &= a_k^2 \varepsilon_k - 2a_k \mathbb{E} [g(p_k)^\top (p_k - p_*)] + b_k \quad (\because a_k \text{は定数}, e_k \perp\!\!\!\perp p_k) \\ &= a_k^2 \varepsilon_k - 2a_k d_k + b_k \\ &\leq (1 - 2\kappa a_k) b_k + a_k^2 \varepsilon_k \quad (\because (11.42) \text{より } d_k \geq \kappa b_k) \end{aligned}$$

が成り立つ. これに, 補題 11.2.4 を適用すると求める結果がしたがう. \square

$g = \nabla L$ の場合は, 不等式 (11.42) は L が強凸関数であり, p_* が L の最小化解であることを意味する. ステップ幅に関する (11.41) はロビンス・モンロー条件 (Robbins-Monro) と呼ばれ, 典型的には $a_k = 1/k$ など, $Ck^{-\alpha}$, $C > 0$, $\alpha \in (0, 1]$ が条件を満たす.

例 11.2.6 (平均値推定と学習率の影響). 実数値確率変数 v の期待値 $p_* := \mathbb{E}[v]$ を求めるために $g(p) := \mathbb{E}[v] - p$ とし, v の独立同分布の標本 v_k が利用可能とする. ここで, v_k を $\mathbb{E}[v]$ の近似値とみなし, $y_k = v_k - p_k$ とする. このとき, y_k

図 11.3: ロビンス・モンロー法 $a_k = Ck^{-\alpha}$.

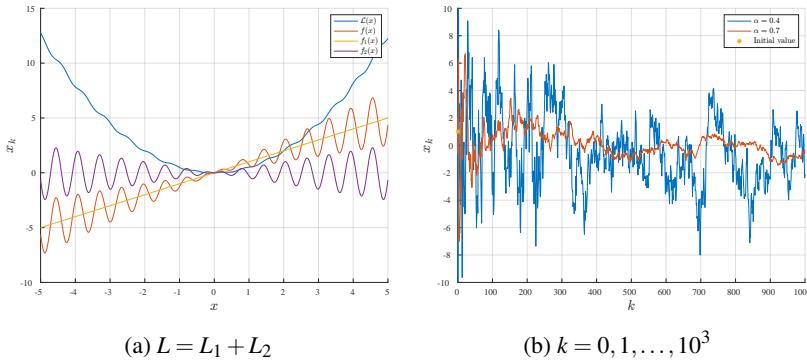
は $g(p_k)$ に $e_k := y_k - g(p_k) = v_k - \mathbb{E}[v_k]$ が付加されているとみなすことができ、この $\{e_k\}$ は平均 0 の独立同分布となる。図 11.3 に $v \sim \mathcal{N}(0, 1)$, $a_k = Ck^{-\alpha}$, $p_0 = 10$ とした場合の p_k を示す。このとき、 α が 1 に近いと a_k は 0 に速く収束し p_* に近づく速度が遅く（図 11.3a），逆に 0 に近いと a_k はなかなか 0 に収束せず p_k の振動が収まらない（図 11.3b）。 \triangleleft

例 11.2.7 (ミニバッチにもとづく確率勾配法)。図 11.4a の L_1 , L_2 および $L := L_1 + L_2$ を考える。ここで、独立同分布の $\iota_k \sim \text{Uni}(\{1, 2\})$ を用いて、 $\nabla L_{\iota_k}(p_k)$ により勾配 ∇L を近似する。このとき、

$$p_{k+1} = p_k - a_k y_k, \quad y_k := \nabla L_{\iota_k}(p_k) \quad (11.45)$$

の p_k は L の最小化解 $p_* = 0$ に収束している（図 11.4b）。 \triangleleft

例 11.2.8 (裾分布の推定)。実数値確率変数 v および $c \in (0, 1)$ に対して、 $\mathbb{P}(v \leq p_*) = c$ となる $p_* \in \mathbb{R}$ を v の標本 $\{v_k\}_k$ ($\{v_k\}$ は独立かつ v と同分布) のみから求める。このとき、 $g(p) := \mathbb{P}(v \leq p) - c$ および $y_k := \mathbb{1}_{(v_k \leq p_k)} - c$ とすると、 $e_k = y_k - g(p_k) = \mathbb{1}_{(v_k \leq p_k)} - \mathbb{P}(v < p_k)$ である。このとき、 $\mathbb{E}[e_k] = 0$ であるが、分布は $p_k \in \sigma(e_{0:k-1})$ に依存し独立でも同分布でもない。 \triangleleft

図 11.4: 確率勾配法 ($p_* = 0$).

この例の場合, $\{e_k\}$ は独立ではないため定理 11.2.5 では扱うことができない. また, 過去の時系列に応じてステップ幅 a_k を調整したい状況もあるだろう. この場合に拡張した結果を付録の定理 B.9 に示す.

注意 11.2.9 (学習率の自動調整). 本節では, 通常の勾配法 $p_{k+1} = p_k - a_k y_k$, $y_k \approx \nabla L(p_k)$ において勾配の値の不確かさがおよぼす影響について解析してきたが, 勾配法そのものの改良の余地がある. たとえば, 代表的なアルゴリズムに

$$p_{k+1} = p_k - av_k, v_{k+1} = \beta v_k + y_k \quad (11.46)$$

により勾配の時間方向に平滑化して更新方向として用いる**モーメンタム法**,

$$p_{k+1} = p_k - a \frac{y_k}{\sqrt{y_k^2 + \epsilon}} m_k, m_{k+1} = \rho m_k + (1 - \rho) y_k^2 \quad (11.47)$$

により勾配を (二乗と平方根は要素ごと) 擬似的に正規化する **RMSProp**, それらの考え方を取り入れた **Adam** などがある. \diamond

12

ガウス過程回帰 (16)

7.1 節において、パラメトリックな関数回帰手法を概説した。本章では、そこで紹介した二次の正則化項を付与したリッジ回帰とカーネルトリックをより一般的な枠組みでとらえなおす形で、ノンパラメトリック手法を紹介する。また、近年、動的システムのモデル化にも広く用いられるガウス過程に関して、その関数解析的な構造と実用的な関数回帰手法を概説する。

12.1 カーネル法

本節の目標は、カーネルトリックにおいて活用されていた性質を抽象化し、**カーネル法**と呼ばれる一連のノンパラメトリック手法の基本的な考え方を理解することである。なお、本章で現れる正規分布は分散が半正定値（正定でない）の退化正規分布である（注意 2.2.8）。

12.1.1 半正定値カーネル

まず、注意 7.1.7 のカーネルトリックの性質を再掲しよう。与えられた特徴写像 $\phi : X \rightarrow \mathbb{R}^{N_f}$ および $p \sim \mathcal{N}(0, I)$ に対して、推定対象の関数 $f_p := \phi^\top p \in \text{rv}(\mathcal{F}_\phi)$ の代入値 $y_s = f_p(x_s) + e_s$, $e_s \sim \mathcal{N}(0, \sigma^2)$ を観測する。ただし、

$$\mathcal{F}_\phi := \{\phi^\top p : p \in \mathbb{R}^{N_f}\} \quad (12.1)$$

である。このとき、データ $\{(x_s, y_s)\}_s$ に対して f_p の MAP 推定関数は

$$\mathcal{F}_{k(\cdot, \mathbf{x}_{1:\bar{s}})} := \{k(\cdot, \mathbf{x}_{1:\bar{s}})^\top \mathbf{p}' : \mathbf{p}' \in \mathbb{R}^{\bar{s}}\} \subset \mathcal{F}_\phi \quad (12.2)$$

の要素であることを示した。ただし、

$$k(\mathbf{x}, \mathbf{y}) := \phi(\mathbf{x})^\top \phi(\mathbf{y}), \mathbf{x}, \mathbf{y} \in X \quad (12.3)$$

$$k(\cdot, \mathbf{x}_{1:\bar{s}}) := \begin{bmatrix} k(\cdot, \mathbf{x}_1) & \dots & k(\cdot, \mathbf{x}_{\bar{s}}) \end{bmatrix}^\top : X \rightarrow \mathbb{R}^{\bar{s}} \quad (12.4)$$

である。なお、 $K_\phi(\mathbf{x}_{1:\bar{s}}) = \phi(\mathbf{x}_{1:\bar{s}})^\top \phi(\mathbf{x}_{1:\bar{s}}) \in \mathbb{R}^{\bar{s} \times \bar{s}}$ の表記のもと、対応する \mathbf{p}' は

$$\sum_s \|y_s - k(\mathbf{x}_s, \mathbf{x}_{1:\bar{s}})^\top \mathbf{p}'\|^2 + \sigma^2 \mathbf{p}'^\top K_\phi(\mathbf{x}_{1:\bar{s}}) \mathbf{p}' \quad (12.5)$$

の最小化解

$$(K_\phi(\mathbf{x}_{1:\bar{s}}) + \sigma^2 I)^{-1} \mathbf{y}_{1:\bar{s}} \quad (12.6)$$

である。

ここで、(12.2)において、(12.3)の形に限定しない一般の二変数関数 k を考えることにしよう。例えば、 $k(\cdot, \mathbf{y})$ が \mathbf{y} にピークをもつ釣鐘形となる

$$k_{SE}(\mathbf{x}, \mathbf{y}) := \exp(-\|\mathbf{x} - \mathbf{y}\|^2/c^2), c > 0 \quad (12.7)$$

のような関数を選べば、 $\phi_s := k(\cdot, \mathbf{x}_s)$, $s = 1, \dots, \bar{s}$ の線形結合からなる(12.2)は「 \mathbf{p}' の選択により、与えられたデータ付近の関数形を局所的に指定する」ことに対応し、直感的にも自然であろう。以下では、この関数 k は $K_\phi(\mathbf{x}_{1:\bar{s}})$ に対応する行列が半正定となることだけを仮定する。

定義 12.1.1 (正定値カーネル)。二変数関数 $k : X \times X \rightarrow \mathbb{R}$ が対称性 $k(\mathbf{x}, \mathbf{x}') = k(\mathbf{x}', \mathbf{x})$ を持ち、任意の N および \mathbf{x}_s , $s = 1, \dots, N$ に対して行列

$$K(\mathbf{x}_{1:N}) := \begin{bmatrix} k(\mathbf{x}_1, \mathbf{x}_1) & k(\mathbf{x}_1, \mathbf{x}_2) & \dots & k(\mathbf{x}_1, \mathbf{x}_N) \\ k(\mathbf{x}_2, \mathbf{x}_1) & k(\mathbf{x}_2, \mathbf{x}_2) & \dots & k(\mathbf{x}_2, \mathbf{x}_N) \\ \vdots & \vdots & \vdots & \vdots \\ k(\mathbf{x}_{\bar{s}}, \mathbf{x}_1) & k(\mathbf{x}_{\bar{s}}, \mathbf{x}_2) & \dots & k(\mathbf{x}_{\bar{s}}, \mathbf{x}_N) \end{bmatrix} \in \mathbb{R}^{N \times N} \quad (12.8)$$

が半正定値行列となるとき, $k(\cdot, \cdot)$ を半正定値カーネル (positive semi-definite kernel), もしくは簡単にカーネルと呼ぶ. \triangleleft

例 12.1.2 (カーネルの例). 代表的な \mathbb{R}^n 上のカーネルに, (12.7) のガウシアンカーネル (Gaussian kernel) があり, SE カーネル (square exponential kernel), RBF カーネル (radially bounded function kernel) とも呼ばれる. また, ラプラスカーネル

$$k(x, y) := \exp(-\alpha \sum_{i=1}^n |x_i - y_i|) \quad (12.9)$$

も半正定値性を満たす. 特に, 釣鐘型の形状をしている必要はなく, 他の例として, ユークリッドカーネル $k(x, y) := x^\top y$, 多項式カーネル $k(x, y) = (c + x^\top y)^d$, $c > 0, d \in \mathbb{N}$ もカーネルの条件を満たす. \triangleleft

もともとは (12.3) で与えられる k を考えていたが, 任意の ϕ に対してこの手順により得られる k もカーネルとなる.

定理 12.1.3 (特徴写像に対応するカーネル). X からある内積空間 H_f への関数 $\phi : X \rightarrow H_f$ に対して,

$$k(x, x') := \langle \phi(x), \phi(x') \rangle_{H_f} \quad (12.10)$$

により定義される k は半正定カーネルである. \triangleleft

証明. 対称性は内積の対称性よりしたがう. また, 任意の $c := [c_1 \ \dots \ c_N]^\top \in \mathbb{R}^N$ に対して,

$$\begin{aligned} c^\top K(x_{1:N})c &= \sum_{i,j=1}^N c_i c_j k(x_i, x_j) = \sum_{i,j=1}^N c_i c_j \langle \phi(x_i), \phi(x_j) \rangle_{H_f} \\ &= \left\langle \sum_{i=1}^N c_i \phi(x_i), \sum_{j=1}^N c_j \phi(x_j) \right\rangle_{H_f} = \left\| \phi(x_{1:N})^\top c \right\|_{H_f}^2 \geq 0 \end{aligned} \quad (12.11)$$

であることより半正定性がしたがう. \square

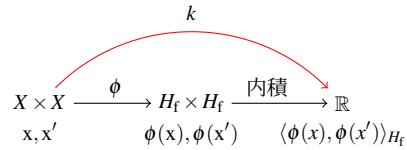


図 12.1: カーネル関数

このカーネルは、 $x, x' \in X$ を非線形関数 ϕ により H_f の要素に変換し、そこでの内積の値 $k(x, x')$ を x, x' の内積であるかのように利用する感覚に近い（図 12.1）。

注意 12.1.4 (マーサー (Mercer) の定理). 逆に適切な仮定と位相のもと、カーネル k は独立な関数列 $\{\phi_i\}_i$ と実数列 λ_i を用いて、

$$k(x, x') = \sum_{i=1}^{\infty} \lambda_i \phi_i(x) \phi_i(x') \quad (12.12)$$

と分解できる [59]. \triangleleft

12.1.2 ガウス過程回帰

$X \subset \mathbb{R}^n$ 上の実数値関数の集合 \mathcal{F} に実現値をもつ確率変数 $f \in \text{rv}(\mathcal{F})$ を設定するにあたり、7.1.1 節では (12.1) の \mathcal{F}_ϕ など特定のパラメータ表現にもとづいて f の事前分布を導入した。これに対して、本節ではノンパラメトリックに事前分布を導入してみよう。具体的には、さまざまな $x_{1:N}$ に対する

$$f(x_{1:N}) = \begin{bmatrix} f(x_1) & \dots & f(x_N) \end{bmatrix}^\top \in \text{rv}(\mathbb{R}^N) \quad (12.13)$$

の統計的性質を規定することで、間接的に f においてどのような関数が実現しやすいかを表現する。

定義 12.1.5 (ガウス過程). 実数値関数 $\mu : X \rightarrow \mathbb{R}$ 、カーネル関数 $k : X \times X \rightarrow \mathbb{R}$ 、 $f : X \rightarrow \text{rv}(\mathbb{R})$ とする。このとき、任意の N および $x_{1:N}$ に対して (12.13) の $f(x_{1:N})$ が

$$f(\mathbf{x}_{1:N}) \sim \mathcal{N}(\mu(\mathbf{x}_{1:N}), K(\mathbf{x}_{1:N})) \quad (12.14)$$

にしたがうとき, f は平均関数 μ , 共分散カーネル k をもつガウス過程 (Gaussian process) であるといい, $f \sim \text{GP}(\mu, k)$ と表記する. ただし, $\mu(\mathbf{x}_{1:\tilde{s}})$ は (12.13) と同様に定義する. \triangleleft

注意 12.1.6 (確率過程). これまでの章では, 時間変数 $k \in \mathbb{Z}_+$ でパラメetrizeされた確率変数の集合を (\mathbb{Z}_+ 上の) 確率過程と呼んできた. より一般的には, このパラメータは離散値をとる必要も, 時間に対応している必要もない. つまり, ガウス過程も \mathbb{R}^n 上の実数値確率過程である. \triangleleft

この定義を満たす \mathbb{R}^n 上の実数値確率過程 f の存在性はまったく自明ではないが, コルモゴロフの拡張定理により保証できる []. こうして, f の関数空間上の事前分布を定めることができた. 任意の $\mathbf{x} \in \mathbb{R}^n$ に対して

$$\mathbb{E}_f[f(\mathbf{x})] := \int_{\mathcal{F}} f(\mathbf{x}) \varphi_f(f) df = \mu(\mathbf{x}) \quad (12.15)$$

が成り立つように, f は平均的に μ に近い関数が実現する傾向がある. また任意の $f \in \mathcal{F}$ および $\mathbf{x}_{1:N}$ に対して, $\varphi_{f(\mathbf{x}_{1:N})}(\mu(\mathbf{x}_{1:N})) \geq \varphi_{f(\mathbf{x}_{1:N})}(f(\mathbf{x}_{1:N}))$ であるという意味でモードを与える関数でもある. 一方で, $k(\mathbf{x}_1, \mathbf{x}_2)$ は確率変数 $f(\mathbf{x}_1)$ と $f(\mathbf{x}_2)$ の共分散である. したがって直観的には, $k(\mathbf{x}_1, \mathbf{x}_2)$ が大きいほど $f(\mathbf{x}_1)$ と $f(\mathbf{x}_2)$ は近い傾向があり, 結果として滑らかな関数が実現しやすい.

例 12.1.7 (ガウス過程). 例えば $\mu(\mathbf{x}) = \mathbf{x}$ に対して $f \sim \text{GP}(\mu, k_{\text{SE}})$ のガウシアンカーネルの例を図 12.2 に示す. この場合, c が大きくなるにつれて滑らかな関数が実現しやすい確率過程となる. \triangleleft

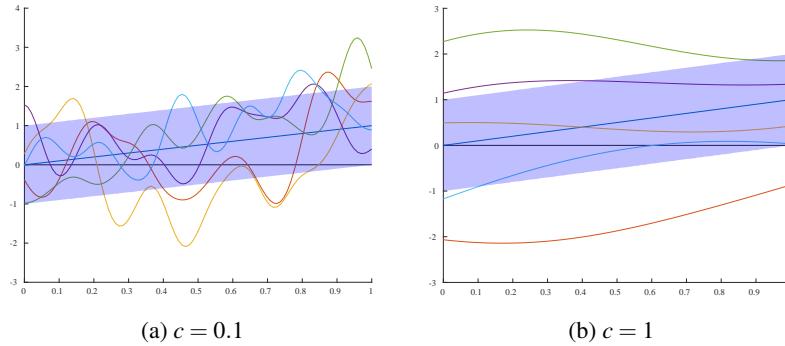


図 12.2: ガウス過程の標本関数. 網掛けは平均関数に標準偏差を加減した領域 $\mu(x) \pm \sqrt{k(x,x)}$.

これで事前分布が設定できたので、これまで同様に特定の x に対して $f(x)$ の雑音付き観測値が得られた場合の、 f の事後分布を求める.

定理 12.1.8 (ガウス過程回帰). f を平均関数 μ および共分散カーネル k をもつガウス過程とする. また、ある \bar{s} および $\mathcal{D} := \{(x_s, y_s)\}_{s=1:\bar{s}}$ に対して

$$y_s := f(x_s) + e_s, \quad e_s \sim \mathcal{N}(0, \sigma^2), \quad s = 1, 2, \dots, \bar{s} \quad (12.16)$$

とし、 f, e_1, \dots は独立であるとする. このとき、 $y_{1:\bar{s}} = y_{1:\bar{s}} \in \mathbb{R}^{\bar{s}}$ のもとでの $f \in \text{rv}(\mathcal{F})$ の事後分布は、平均関数および共分散カーネル関数が

$$\mu(\bar{x}|\mathcal{D}) := \mu(\bar{x}) + k(\bar{x}, x_{1:\bar{s}})^\top (K(x_{1:\bar{s}}) + \sigma^2 I)^{-1} (y_{1:\bar{s}} - \mu(x_{1:\bar{s}})) \quad (12.17)$$

$$k(\bar{x}, \bar{x}'|\mathcal{D}) := k(\bar{x}, \bar{x}') - k(\bar{x}, x_{1:\bar{s}})^\top (K(x_{1:\bar{s}}) + \sigma^2 I)^{-1} k(\bar{x}', x_{1:\bar{s}}) \quad (12.18)$$

で与えられるガウス過程である. \diamond

証明. $N \in \mathbb{Z}_+$ および $x_i^* \in X, i = 1, 2, \dots, N$ を任意にとる. このとき、 $x_\cdot := x_{1:\bar{s}}, x_\cdot^* := x_{1:N}^*, y_\cdot := y_{1:\bar{s}}$ と表記すると、ガウス過程の定義より、

$$\begin{bmatrix} y \\ f(\mathbf{x}_:) \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} \mu(\mathbf{x}_:) \\ \mu(\mathbf{x}_:) \end{bmatrix}, \begin{bmatrix} K(\mathbf{x}_:) + \sigma^2 I & k(\mathbf{x}_:^\ast, \mathbf{x}_:) \\ k(\mathbf{x}_:^\ast, \mathbf{x}_:)^\top & K(\mathbf{x}_:) \end{bmatrix} \right) \quad (12.19)$$

$$k(\mathbf{x}_:^\ast, \mathbf{x}_:) := \begin{bmatrix} k(\mathbf{x}_1^\ast, \mathbf{x}_:) & \dots & k(\mathbf{x}_N^\ast, \mathbf{x}_:) \end{bmatrix} \in \mathbb{R}^{\bar{s} \times N} \quad (12.20)$$

である。したがって、定理 2.3.5 より、 $f(\mathbf{x}_:^\ast) \in \text{rv}(\mathbb{R}^N)$ の事後分布は

$$f(\mathbf{x}_:^\ast) |_{y_:=y_:} \sim \mathcal{N} \left(\mu(\mathbf{x}_:^\ast | \mathcal{D}), K(\mathbf{x}_:^\ast) - k(\mathbf{x}_:^\ast, \mathbf{x}_:)^\top (K(\mathbf{x}_:) + \sigma^2 I)^{-1} k(\mathbf{x}_:^\ast, \mathbf{x}_:) \right) \quad (12.21)$$

で与えられる。この共分散行列の (i, j) 成分は、

$$k(\mathbf{x}_i^\ast, \mathbf{x}_j^\ast) - k(\mathbf{x}_i^\ast, \mathbf{x}_:)^\top (K(\mathbf{x}_:) + \sigma^2 I)^{-1} k(\mathbf{x}_j^\ast, \mathbf{x}_:) \quad (12.22)$$

であることから、求める結果が従う。 \square

これによりガウス過程に対して、任意の評価点における観測値が得られたとき、事後確率過程もガウス過程となることが示された。ここで、事後共分散カーネル $k(\cdot, \cdot | \mathcal{D})$ は y に依存しない。したがって、各点における事後分散の大きさは観測がおこなわれた場所のみにより決まり、「観測値と真の値が近い箇所では事後分散が小さい」といった一見もっともらしい解釈は正しくない。

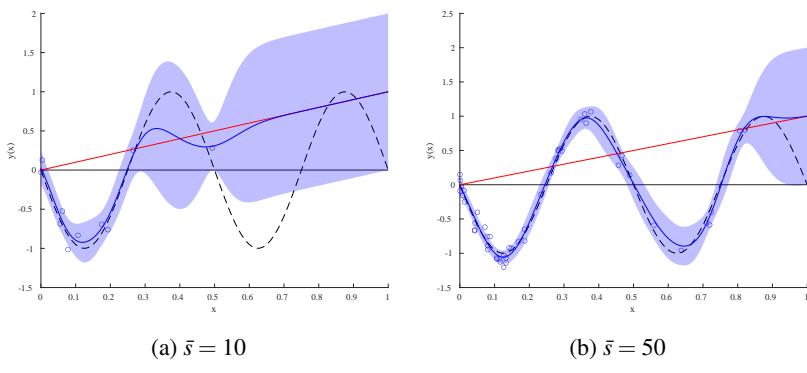


図 12.3: RBF カーネルによるガウス過程回帰。○は $\{(\mathbf{x}_s, \mathbf{y}_s)\}_{s=1:\bar{s}}$

194 12. ガウス過程回帰 (16)

例 12.1.9. 例 12.1.7 と同じガウス過程 f を考える ($\sigma = 0.1$). ここで, x_s はランダムに生成し,

$$y_s = -\sin(4\pi x_s) + e_s, \quad e_s \sim \mathcal{N}(0, 0.1^2) \quad (12.23)$$

の観測値を用いたガウス過程回帰の結果を図 12.3 に示す. 観測点が多い領域ほど分散が小さく, 観測点が少ない領域では平均値関数が事前平均関数に近くなることがわかる. \diamond

最後に, ガウス過程回帰における MAP 推定関数に相当する平均値関数 $\mu(\cdot|\mathcal{D})$ を, パラメトリック推定の場合と比較する. 事前平均 μ からの差 $f - \mu$ を新たに f とおくことで一般性を失わず $\mu = 0$ とする. このとき,

$$\mu(\cdot|\mathcal{D}) = k(\cdot, x_{1:\bar{s}})^\top (K(x_{1:\bar{s}}) + \sigma^2 I)^{-1} y_{1:\bar{s}} \quad (12.24)$$

であるが, これは (12.6) で得られる回帰関数に他ならない. つまり, ノンパラメトリック回帰として定式化したものの, 良くも悪くもパラメトリック回帰と同じ結果が得られたことになる. したがって, 事後分散関数により近似誤差が理論的に評価できる点, (12.2) の単純な関数形に限定する正当性を与えており, カーネルさえ選択すれば系統的かつ柔軟な手続きが与えられる点こそが大きな利点と言える.

計算量の観点からは, カーネル行列の逆行列の計算がデータ数 \bar{s} とともに増大する点が課題である. たとえば, 図 12.3 では, $x = 0.1$ 付近のデータは間引いても精度はそれほど変わらないことが予想される. こうした代表点（誘導点と呼ばれる）の選定方法に関しても様々な研究がある.

演習 12.1. 図 12.2 は $x_i =$ として, (有限次元の) 正規分布 $\mathcal{N}(\mu(x_{1:N}), K(x_{1:N}))$ の標本を生成し, 直線補間してプロットした結果である. 同様の手順により, 図 12.3 の事後分布にしたがう確率過程の標本関数を計算機を用いて再現せよ. \diamond

12.1.3 【オプション】再生核ヒルベルト空間

定理 7.1.2 において, パラメトリック回帰において統計推論による定式化と

最小二乗法としての定式化の等価性を確認した。前節ではノンパラメトリック回帰に対する統計推論による定式化を紹介したので、本節では最小二乗法としての定式化を検討する。これには、関数形を限定することなく、適切な滑らかさを有する関数の集合を導入しなければならない。そこで、つぎの性質をもつヒルベルト空間に注目する。

定義 12.1.10 (再生核ヒルベルト空間)。 \mathcal{H} を X 上の実数値関数を要素とし、内積 $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ を有するヒルベルト空間とする。このとき、任意の $x \in X$ に対して、ある $C_x < \infty$ が存在し

$$|f(x)| \leq C_x \|f\|_{\mathcal{H}}, \forall f \in \mathcal{H} \quad (12.25)$$

が成り立つとき、 \mathcal{H} は X -上の再生核ヒルベルト空間 (Reproducing Kernel Hilbert Space; RKHS) であるという。 ◇

定理 12.1.3 では \mathbb{R}^{N_f} の一般化としての内積空間 H_f を導入したが、このヒルベルト空間 \mathcal{H} は \mathcal{F}_ϕ に対応する概念である。どちらにおいても内積が重要な役割を果たすが、混同しないように注意してほしい。ここで定義された再生核ヒルベルト空間は、「代入」という操作 $f \mapsto f(x)$ が有界 ($\sup_x C_x < \infty$ は不要) という意味で、なめらかな関数からなる関数族である。関数を要素としてもつヒルベルト空間の代表例である $\langle f, f' \rangle := \int f(x)f'(x)dx$ をもつ L^2 -空間がこの性質をもたないことを確認しておく。

例 12.1.11 (L^2 における代入操作の非有界性)。任意の $c > 0$ に対して関数 $f_c(x) := \exp(-\|x\|^2/c^2)$ は L^2 の要素であるが、 $x = 0$ を代入するという操作は $f_c(0) = 1, \forall c$ であるのに対して、 $c \rightarrow 0$ のとき $\|f_c\|_{L^2} \rightarrow 0$ である。このように、「いくらでも急峻に変動する関数」を含むような集合において「急峻な変動を加味しないノルム」を導入すると、代入操作は有界とならない。関連して、 $g_\sigma(x) := \mathcal{N}(x|0, \sigma^2) \in L^2$ を考えると、十分なめらかな任意の関数 f に対して、 $\sigma \rightarrow 0$ の極限において $\langle f, g_\sigma \rangle \rightarrow f(0)$ が成り立つ (注意 2.2.8)。また、コーシー・シュワ

196 12. ガウス過程回帰 (16)

ルツの不等式より $\langle f, g_\sigma \rangle \leq \|f\|_{L^2} \|g_\sigma\|_{L^2}$ も成り立つ。しかしながら, g_σ はディラックのデルタ関数に収束し (注意 2.2.8), $\sigma \rightarrow 0$ で L^2 にとどまらない。△

再生核ヒルベルト空間はとても自由度の高い集合にみえるが, 実際には対応する唯一のカーネル関数により特徴づけられる。

定理 12.1.12 (再生核ヒルベルト空間とカーネルの対応). 任意の X -上の再生核ヒルベルト空間 \mathcal{H} に対して, 二つの条件

- 任意の $x \in X$ に対して, $k(\cdot, x) \in \mathcal{H}$
- (再生性) 任意の $f \in \mathcal{H}$, $x \in X$ に対して,

$$\langle f(\cdot), k(\cdot, x) \rangle_{\mathcal{H}} = f(x) \quad (12.26)$$

を満たす唯一の X -上のカーネル k が存在し, これを \mathcal{H} の**再生核**と呼ぶ。逆に, 任意の半正定値カーネル関数 k に対して, 再生核 k をもつ再生核ヒルベルト空間 \mathcal{H}_k が唯一つ存在する。△

証明. リースの表現定理 [60] により, ヒルベルト空間上の有界線形汎関数は内積により表現される。したがって, \mathcal{H} の線形汎関数 $f \mapsto f(x)$ が有界線形汎関数であるため, 任意の $x \in X$ に対してある $k(\cdot, x) \in \mathcal{H}$ が存在して, (12.26) が成り立つ。内積の対称性から,

$$k(x, y) = \langle k(\cdot, x), k(\cdot, y) \rangle = \langle k(\cdot, y), k(\cdot, x) \rangle = k(y, x) \quad (12.27)$$

である。半正定値性は, (12.11) と同様に,

$$\begin{aligned} c^\top K(x_{1:N})c &= \sum_{i,j=1}^n c_i c_j k(x_i, x_j) = \sum_{i,j=1}^n c_i c_j \langle k(\cdot, x_i), k(\cdot, x_j) \rangle_{\mathcal{H}} \\ &= \left\langle \sum_{i=1}^n c_i k(\cdot, x_i), \sum_{j=1}^n c_j k(\cdot, x_j) \right\rangle_{\mathcal{H}} = \left\| k(\cdot, x_{1:\bar{s}})^\top c \right\|_{\mathcal{H}}^2 \geq 0 \end{aligned} \quad (12.28)$$

であることよりしたがう。

後半の \mathcal{H}_k はすべての $\{k(\cdot, x)\}_{x \in X}$ を含むベクトル空間を完備化することで得られる。ムーア・アロンシャイン (Moore-Aronszajn) の定理として知られる一意性は省略する [61]。□

再生性は代入操作がある関数との内積として表現できることを意味しているが、具体例をみてみよう。

例 12.1.13 (ReLU 関数と RKHS). \mathbb{R}_+ 上で定義された絶対連続かつ $f(0) = 0$ を満たす関数からなる集合に内積

$$\langle f, g \rangle_{\mathcal{H}} := \int_0^\infty \frac{df}{dx}(x) \frac{dg}{dx}(x) dx \quad (12.29)$$

を導入したヒルベルト空間 \mathcal{H} を考える。したがって、

$$\|f\|_{\mathcal{H}}^2 := \int_0^\infty \left| \frac{df}{dx}(x) \right|^2 dx \quad (12.30)$$

である。このとき、この \mathcal{H} は再生核

$$k(x, x') := \min(x, x') = -\text{ReLU}(-(x - x')) + x' \quad (12.31)$$

をもつ再生核ヒルベルト空間となる。実際、 $k(x, x')$ の x に関する一階微分は $\mathbb{1}_{\{c < x'\}}(x)$ 、二階微分は $-\delta_x(x)$ に注意すると、(12.29) の部分積分から

$$\langle f, k(\cdot, x') \rangle_{\mathcal{H}} = [f \cdot \mathbb{1}_{\{c < x'\}}]_0^\infty - \int_0^\infty f \cdot (-\delta_{x'}(x)) dx = f(x') \quad (12.32)$$

が確かめられる。 \triangleleft

これらの準備のもと、最小二乗法による解釈は次のように一般化される。

定理 12.1.14 (リプレゼンター定理). 再生核 k をもつ X 上の再生核ヒルベルト空間 \mathcal{H}_k を考える。このとき、 $\mathcal{D} := \{(x_s, y_s)\}_{s=1}^{\bar{s}}$ および $L: (X \times \mathbb{R}^2)^{\bar{s}} \rightarrow \mathbb{R} \cup \{\infty\}$ 、厳密に単調増加する関数 $r: [0, \infty) \rightarrow \mathbb{R}$ 、 \mathcal{D} が与えられたとき、

$$L((x_1, y_1, f(x_1)), \dots, (x_{\bar{s}}, y_{\bar{s}}, f(x_{\bar{s}}))) + r(\|f\|_{\mathcal{H}_k}) \quad (12.33)$$

を最小化する $f \in \mathcal{H}_k$ は (12.2) の $\mathcal{F}_{k(\cdot, x_1, \bar{s})}$ の要素である。 \triangleleft

証明. 表記の簡単化のため $\mathcal{F}_{k(\cdot, x_1, \bar{s})}$ を \mathcal{F} とあらわす。以下では、ヒルベルト空間における標準的な射影定理を利用する。まず、 \mathcal{H}_k における \mathcal{F} の直交補

198 12. ガウス過程回帰 (16)

空間 $\mathcal{F}^\perp := \{f \in \mathcal{H}_k : \langle f, g \rangle = 0, \forall g \in \mathcal{F}\}$ を導入する。任意の $f \in \mathcal{H}_k$ は、 $f = f_+ + f_\perp$, $f_+ \in \mathcal{F}$, $f_\perp \in \mathcal{F}^\perp$ と分解でき、 $\|f\|^2 = \|f_+\|^2 + \|f_\perp\|^2$ である。ここで、 $\phi_s := k(\cdot, x_s) \in \mathcal{F}$ より、 $f_\perp(x_s) = \langle f, \phi_s \rangle = 0, \forall s$ に注意すると、 f_\perp は (12.33) の第一項に影響を与えない。ゆえに、 $f_\perp = 0$ とすることが最適である。□

系 12.1.15 (ガウス過程回帰の事後平均との対応)。定理 12.1.14 と同じ \mathcal{H}_k および \mathcal{D} に対して、

$$\sum_{s=1}^{\bar{s}} |f(x_s) - y_s|^2 + \sigma^2 \|f\|_{\mathcal{H}_k}^2 \quad (12.34)$$

の最小化する $f \in \mathcal{H}_k$ の一つは (12.24) の $\mu(\cdot | \mathcal{D}) \in \mathcal{F}_{k(\cdot, x_{1:\bar{s}})}$ である。△

証明。 定理 12.1.14 より求める最小化解は $\mathcal{F}_{k(\cdot, x_{1:\bar{s}})}$ の要素であるため、 $c \in \mathbb{R}^{\bar{s}}$ を用いて $k(\cdot, x_{1:\bar{s}})^\top c$ とかける。また、(12.28) から $\|k(\cdot, x_{1:\bar{s}})^\top c\|_{\mathcal{H}_k}^2 = c^\top K(x_{1:\bar{s}})c$ が成り立つ。したがって、定理 7.1.5 と同様に求める結果がしたがう。□

例えば、例 12.1.13 の RKHS の場合、近似誤差に (12.30) を加えた値を最小化する意味で滑らかな関数を求めていることに対応する。この定理より、最適な関数は評価点のみに節点をもつ区分的線形関数であることがわかる（図 12.4）。繰り返しになるが、例 12.1.13 の RKHS は非常に一般的な関数を含むが、得られる MAP 推定に対応する事後平均関数は単純な区分線形関数である。

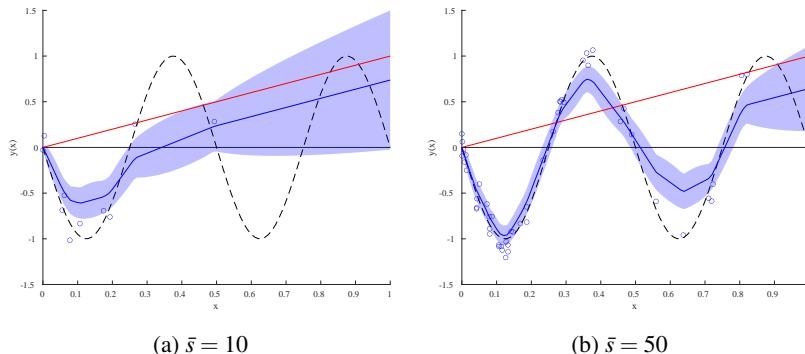


図 12.4: $k(x, y) = \min(x, y)$ によるガウス過程回帰。○は $(x_i, y_i)_{i=1:N}$

12.2 ガウス過程回帰状態空間モデル

前章までは、 $\mathbb{X} \subset \mathbb{R}^{n_x}$, $\mathbb{U} \subset \mathbb{R}^{n_u}$ とし、雑音 v と（確定的な）関数 f を用いて $x_{k+1} = f(x_k, u_k, v_k)$ により確率的な状態遷移を表現していた。ここでは、確率性を関数に押し込み、 $\mathbb{X} \times \mathbb{U}$ から \mathbb{X} への関数値の確率変数 f （すなわち確率過程 $f: \mathbb{X} \times \mathbb{U} \rightarrow \text{rv}(\mathbb{X})$ ）を用いて、

$$x_{k+1} = f(x_k, u_k) \quad (12.35)$$

により状態遷移の不確実性を表現する。特に状態遷移関数をガウス過程回帰により与えるモデルをガウス過程状態空間モデルと呼ぶ。本書では、実数値をとるガウス過程しか考えていないため、 f の i 番目の要素 $f_i: \mathbb{R}^{n_x+n_u} \rightarrow \mathbb{R}$ が平均関数 $\mu_i(\cdot)$ 、共分散カーネル $k_i(\cdot, \cdot)$ をもつ $\mathbb{X} \times \mathbb{U}$ -上のガウス過程回帰とする。もしも、状態遷移関数をあらわす確率過程 f_i が各要素および各時刻で独立であるとすると、

$$x_{k+1} = \mu(x_k, u_k) + \text{diag}\{\sqrt{k_i((x_k, u_k), (x_k, u_k))}\}v_k, \quad v_k \sim \mathcal{N}(0, I) \quad (12.36)$$

$$\mu := \begin{bmatrix} \mu_1 & \cdots & \mu_n \end{bmatrix}^\top \quad (12.37)$$

となり平均値関数をノミナルモデルとし、状態依存雑音が加法的に加わると見なせる。

つぎに、状態・入力の組み (x_k, u_k) の標本系列 x_i, u_i が利用可能であると仮定し、 f_i の推定方法を考える。 (12.35) であることに注意すると、カーネルを適当に定め、 $k = 1, 2, \dots$ に対して $\tilde{x}_k := \begin{bmatrix} x_k^\top & u_k^\top \end{bmatrix}^\top$ を入力、 x_{k+1} の第 i 成分を出力データとして扱えば、 $\mathcal{D} := \{(\tilde{x}_k, (x_{k+1})_i)\}$ を定理 12.1.8 に適用し事後平均関数 $\mu_i(\cdot | \mathcal{D})$ と事後共分散カーネル $k_i(\cdot, \cdot | \mathcal{D})$ をもつガウス過程が得られる。もしも事前に粗いモデルなどが得られているのであれば、それらを事前平均に用いることも有用であり、その場合はモデル化誤差をデータから回帰することに相当する。

200 12. ガウス過程回帰 (16)

以下では、12.2.1 節においてガウス過程の引数が確率変数である場合の平均と分散を解析的に計算する手法を紹介し、ガウス過程状態空間モデルの時間発展の計算に応用する。また、12.2.2 節ではベイズ最適化と呼ばれる最適化手法を紹介し、能動的なシステム同定用データ選択に適用する。

12.2.1 モーメントマッチング

(1) 正規分布のガウス過程への代入 一般に、ガウス過程 $f \sim \text{GP}(\mu, k)$ と $x \in X$ に対して $f(x) \in \text{rv}(\mathbb{R})$ であったが、 f と独立な $x \in \text{rv}(X)$ に対して $f(x) \in \text{rv}(\mathbb{R})$ の性質を調べる。確率過程 f は確率的関数が定まるが、その代入値にも不確かさがある場合の出力に相当する。

ここで、再びガウス関数の特殊な性質を活用するために、 $\alpha \in \mathbb{R}^N$ 、正定行列 $M \in \mathbb{R}^{N \times N}$ および (12.7) のガウシアンカーネル k_{SE} を用いて定義される

$$\mu(\cdot) := k_{\text{SE}}(\cdot, x_{1:\bar{s}})^\top \alpha, \quad (12.38)$$

$$k(x, y) := k_{\text{SE}}(x, y) - k_{\text{SE}}(x, x_{1:\bar{s}})^\top M^{-1} k_{\text{SE}}(y, x_{1:\bar{s}}) \quad (12.39)$$

を平均関数および共分散カーネルとするガウス過程を考えよう。さらに $x \sim \mathcal{N}(\mu_x, \Sigma_x)$ は f と独立であると仮定する。

まず期待値を計算すると塔特性より

$$\mathbb{E}[f(x)] = \int \varphi_x(x) \mathbb{E}[f(x)] dx = \int \varphi_x(x) k_{\text{SE}}(x, x_{1:\bar{s}})^\top dx \alpha \quad (12.40)$$

が成り立つ。ここで被積分関数の第 j 成分は $\mathcal{N}(x | \mu_x, \Sigma_x) \exp(-\|x - x_j\|^2/c^2)$ であるが、ガウス関数同士の積はやはりガウス関数となるため、補題??と同様の計算により積分は陽に書き下せる。また、分散に関しても

$$\mathbb{E}[f(x)^2] = \int \varphi_x(x) \mathbb{E}[f(x)^2] dx \quad (12.41)$$

$$= \int \varphi_x(x) (\mu(x)^2 + k(x, x)) dx \quad (12.42)$$

となる。これに関しても、煩雑にはなるがガウス関数の積が繰り返し現れるだけである。最後に、 f_i , $i = 1, 2$ が平均関数 $\mu_i(x) := k_{\text{SE}}(x, x_{1:\bar{s}})^\top \alpha_i$ をもつガウ

ス過程であるとする。ここで、 f_1, f_2, x が独立ならば、

$$\mathbb{E}[f_1(x)f_2(x)] = \int \varphi_x(x) \mathbb{E}[f_1(x)f_2(x)] dx \quad (12.43)$$

$$= \int \varphi_x(x) \mu_1(x) \mu_2(x) dx \quad (12.44)$$

が成り立つ。これも同様にガウス関数の積分であり、具体的な表現が得られる。

(2) 予測分布の近似計算 (12.35) によって標本経路を発生することはできるが、線形システムの状態推定のように状態変数の分布の時間発展が効率よく計算できれば、大きな利点がある。そこで、非線形フィルタリングの場合のように x_k, u_k の分布を正規分布により近似する。たとえば、 $u_k = K_k x_k + \beta_k$ のように状態の線形フィードバック制御則を考えるとき、 x_k が正規分布にしたがうならば、 $\tilde{x}_k = \begin{bmatrix} x_k^\top & u_k^\top \end{bmatrix}^\top$ も（退化）正規分布にしたがう。したがって、前項の議論より x_{k+1} の平均と分散は解析的に計算することができる。また、 x_{k+1} をこの平均と分散をもつ正規分布により近似（モーメントマッチングと呼ばれる）すれば、逐次的に分布の時間発展を書き下すことができる。

また、

$$\ell_k(x) := 1 - \exp\left(-\frac{1}{2\sigma_\ell^2} \|x_* - x\|^2\right) \quad (12.45)$$

のような状態コストを考えると、 $\mathbb{E}[\ell_k(x_k)]$ も解析的に計算できるため、長期コスト

$$\mathbb{E}\left[\sum_k \ell_k(x_k)\right] \quad (12.46)$$

の K_k や β_k に関する偏微分も解析的な近似計算が可能であり、勾配法により最適化できる。線形の制御則以外にも、決定論的ガウス過程方策

$$u_k = \pi(x_k) := k_{SE}(x_k, x_{1:N_\pi}^\pi)^\top p_\pi, \quad x_i^\pi \in X, \quad i = 1, 2, \dots, N_\pi \quad (12.47)$$

なども、ある程度同様の議論が成立する。本節では計算の方針のみを紹介したのみであるが、詳細や具体的な公式は [62] を参照してほしい。

12.2.2 ベイズ最適化

関数 $f: X \rightarrow \mathbb{R}$ に対して, $x^* := \arg \max_{x \in X} f(x)$ を求める最適化問題を考える。ただし, f の具体的な関数形は未知であり, 指定した $x_s \in X$ に対して

$$y_s := f(x_s) + e_s \quad (12.48)$$

のみが観測できるものとする。こうした問題はブラックボックス最適化と呼ばれる。関数回帰同様, f に関して何らかの滑らかさの仮定をおかなければ, 意味のある理論解析是不可能である。

そこで, この未知関数をガウス過程 $f \sim GP(\mu, k)$ として定式化する。与えられた (x_s, y_s) に対して, f の事後分布 $f|_{\mathcal{D}} \sim GP(\mu(\cdot|\mathcal{D}), k(\cdot, \cdot|\mathcal{D}))$ は定理 12.1.8 により容易に計算できる。ここで新しく検討するのはつぎの評価点 x_{s+1} の選択方法である。関数 f に関して不確実ながら情報がある状況では, 評価点の選択に工夫の余地がある。例えば, s 番目までの最適値 $\bar{f}_s := \arg \max_{i \in [1:s]} y_i$ に対して最大値の改善する確率 (probability of improvement)

$$PI(x) := \mathbb{P}(f(x|\mathcal{D}) > \bar{f}_s) \quad (12.49)$$

または改善量の期待値 (expected improvement)

$$EI(x) := \text{ReLU}(f(x|\mathcal{D}) - \bar{f}_s) \quad (12.50)$$

を最大化する $x \in X$ を x_{s+1} とすることは合理的であろう。もしくは, 重み定数 $\beta > 0$ を用いて平均と分散の和 (upper confidence bound; UCB)

$$UCB(x) := \mu(x|\mathcal{D}) + ck(x, x|\mathcal{D}) \quad (12.51)$$

による MAP 推定関数を活用した最大化と未知領域の探索のバランスをとるべきかもしれない。こうした事後分布を活用したブラックボックス最適化手法を一般的にベイズ最適化 (Bayesian optimization) と呼ぶ。獲得関数 (acquisition function) と呼ばれる (12.49) や (12.51) などの指標の妥当性は問題の性質に依存し,もちろん獲得関数の最大化そのものに必要な計算量も無視できないが, ベイズ最適化が有効な手法となる例も多く報告されている。

演習 12.2 (ベイズ最適化). 図 12.4 において, (12.51) にもとづいてつぎの観測点をきめるとき, c の選び方が選択結果に与える影響を考察せよ. \triangleleft

状態遷移関数の回帰においても, ベイズ最適化の視点は有用である. 例えば, 事後分散の高い状態や入力に対する代入値のデータが得られると, それだけモデルの精度が向上することが期待できる. 実際, 前節のモーメントマッチングにもとづく制御則の最適化とベイズ最適化にもとづく回帰情報の取得を繰り返しおこなう手法が [62] にまとめられている. また PILCO (Probabilistic Inference for Learning Control) というプロジェクトに関する著者のウェブページにおいてソースコードが公開されており, 手軽に試してみることができる.

13|| おわりに

本書では、確率システム論に関して、システム制御理論と統計的学習理論の橋渡しとなるであろう話題を概説してきた。「はじめに」で述べた理由により、一貫して離散時間システムを取り扱ったが、3, 4, 5, 6 章の緒結果には、ほぼそのまま連続時間に対応した結果がある [63, 64]。たとえば、連続時間 LQR 問題に対しては確実性等価原理が成立し、制御則と状態推定器は分離設計できる。また、観測値系列に条件付けられた状態の事後分布の時間発展はクシュナー (Kushner) 方程式、ザカイ (Zakai) 方程式と呼ばれる偏微分方程式で書き下せる。他にも、4.2.3 節の随伴変数にもとづく手法は連続時間制御系設計にも有効であるが、同様に考え方にもとづく「連続的な層」をもつニューラルネットワーク (Neural Ordinary Differential Equation) は機械学習分野でもよく知られている。

もちろん、通信容量制約下の制御と情報理論の関係 [] やランダマイズドアルゴリズム [] など紙幅の都合上、紹介できなかった話題も多々ある。その中でも、5.3 節で議論した双対性や 10.1.2 節で触れたカルバック・ライブラ一制御に続く一連の研究は、特に重要であると考えている。たとえば、エントロピー正則化により意図的に確率性を導入する制御則が、モデル化誤差に対するロバスト性を獲得するだけでなく、Control as Inference という制御を推論として捉える新しい視点を提供する、といった事実は理論的にも実用的にも大変興味深い。また、最適輸送理論と呼ばれる確率分布間の距離に関する理論も確率制御と密接に関連する。こうした話題は、確率システムの面白さや美しさが端的にあらわれるだけでなく、近年、機械学習、ロボティクス、統計力学、脳科学との接点となりつつあるので、機会があれば改めて系統的に紹介したい。

付 錄

A 測度論的確率論

A.1 確率と確率変数

本節では、1.2節で紹介した確率に関する考え方を、測度論の概念を用いて定式化する。1.2節の前提のあとに主観的イメージを述べたが、「 $v_i \in \text{rv}(V_i)$, $i = 1, 2, \dots$ がすべての不確実性の発生源であり、 $\Omega := V_1 \times V_2 \times \dots$ に実現値をもつ $v := (v_1, v_2, \dots) \in \text{rv}(\Omega)$ の結合確率密度関数 $\varphi_v(\omega)$, $\omega \in \Omega$ が唯一つ準備されており、他の確率変数はすべて v の関数 ($\sigma(v)$ の要素) である」とすれば、本節の内容に随分と近づく。大きな違いは、発生源の数が有限個や加算個でなくともよい点と、基本的に確率密度関数 φ_v に頼らず確率 \mathbb{P}_v のみで舞台を構築していく点である。ここで、 \mathbb{P}_v の引数は集合であることから、必然的に集合論が中心的な役割を果たす。

(1) 確率空間 上記の Ω および \mathbb{P}_v に相当するものを導入するにあたり、それらに本質的に要請されている性質を抽出するとつぎに定義にいたる。

定義 A.1 (可測空間・測度空間). 集合 Ω の部分集合を要素とする集合族 \mathcal{E} が

- 1) $\Omega \in \mathcal{E}$.
- 2) $A \in \mathcal{E}$ のとき, $A^c = \Omega \setminus A \in \mathcal{E}$.
- 3) $A_j \in \mathcal{E}$, $j = 1, 2, \dots$ ならば $\bigcup_{j=1}^{\infty} A_j \in \mathcal{E}$.

を満たすとき \mathcal{E} を σ -加法族、 \mathcal{E} の元を可測集合、 Ω と \mathcal{E} の組 (Ω, \mathcal{E}) を可測空間と呼ぶ。さらに、

- 1) $\mu(\emptyset) = 0$.

206 13. おわりに

- 2) $E_j \in \mathcal{E}, j = 1, 2, \dots$ が互いに素, つまり, $E_i \cap E_j = \emptyset (1 \leq i < j < \infty)$ な
らば,

$$\mu \left(\bigcup_{j=1}^{\infty} E_j \right) = \sum_{j=1}^{\infty} \mu(E_j).$$

を満たす集合関数 $\mu : \mathcal{E} \rightarrow [0, \infty]$ を測度と呼び, 可測空間 (Ω, \mathcal{E}) に測度 μ を付
与した 3 つ組 $(\Omega, \mathcal{E}, \mu)$ を測度空間と呼ぶ. \triangleleft

とくに条件 2) は測度 μ の σ -加法性と呼ばれる. 測度は $[0, \infty]$ に値をとるが,
負の値も許し値域を $[-\infty, \infty]$ に拡張した μ を符号付き測度と呼ぶ. つぎに, Ω
を定義域とする関数にも, 可測性の概念を導入する.

定義 A.2 (可測関数). $(\Omega, \mathcal{E}), (X, \mathcal{X})$ をそれぞれ可測空間とする. 関数 $x : \Omega \rightarrow X$
が (\mathcal{E} -) 可測関数であるとは, 任意の $B \in \mathcal{X}$ に対して, $x^{-1}(B) := \{\omega : x(\omega) \in B\} \in \mathcal{E}$ が成り立つことをいう. \triangleleft

測度空間 $(\Omega, \mathcal{E}, \mu)$ を考えると, この定義における x の可測性は任意の $B \in \mathcal{X}$
の逆像の大きさが μ で測れる ($\mu(x^{-1}(B))$ が定義されている) ことを意味する.
ここまで用意した測度論の言葉を用いると, 確率を議論する舞台がつぎのよ
うに準備できる.

定義 A.3 (確率空間). 可測空間 (Ω, \mathcal{E}) 上の測度 \mathbb{P} が

$$\mathbb{P}(\Omega) = 1$$

を満たすとき, \mathbb{P} を確率測度と呼ぶ. 可測空間に確率測度を付与した測度空間
 $(\Omega, \mathcal{E}, \mathbb{P})$ をとくに確率空間, 確率空間の Ω を標本空間, \mathcal{E} の要素を事象とそれ
ぞれ呼ぶ. さらに, (X, \mathcal{X}) を可測空間としたとき, 可測関数 $x : \Omega \rightarrow X$ を X -
値確率変数と呼び, $x \in \text{rv}(X)$ と表記する. \triangleleft

定義 A.3において, すべての確率変数はただ一つの標本空間 Ω 上に, 確率測
度 \mathbb{P} とは無関係に定義されている. 以下では, 確率変数の等号 $x = y$ は $\mathbb{P}(\{\omega :
x(\omega) = y(\omega)\}) = 1$ であることを意味し, \mathbb{P} のもとでほとんど確かに (almost

surely) 等しい (\mathbb{P} -a.s.) とも表記される (不等号などについても同様). また, $x \in \text{rv}(X)$ および関数 $f: X \rightarrow Y$ に対して, 合成関数 $\omega \mapsto f(x(\omega))$ の意味で $f(x) \in \text{rv}(Y)$ も定義する. 確率空間は Ω の要素に \mathbb{P} により重み付けをしているのみで, ランダムさとは何かといった問い合わせに積極的に答えるものではない. 本書では, 環境雑音, 情報の不完全性, モデル化誤差など理工学においてあらわれる様々な不確実性を, 『 Ω 上の重み』として統一的に扱う. ここで, $\Omega, \mathcal{E}, \mathbb{P}$ およびすべての確率変数は, 具体的に書き下せる必要はないが完全に既知, すなわち不確実性を含まないことに注意する.

つぎに, 主に実数値をとる確率変数に興味があるため, 直線上の「長さ」を一般化することにより, 可測空間としての \mathbb{R} の性質を整理しておこう.

定義 A.4 (ボレル集合族・ルベーグ測度[†]). 任意の $a < b$ に対する区間 $[a, b)$ をふくむ最小の σ -加法族が一意に存在し, これをボレル集合族 $\mathcal{B}(\mathbb{R})$ と呼ぶ. さらに, 任意の $a < b$ に対して, $\mu([a, b)) = b - a$ が成り立つ測度 μ が一意に存在し, これをルベーグ測度 μ_L と呼ぶ. \triangleleft

可測空間 $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ 上のルベーグ測度 μ_L を用いれば, 区間だけでなく, 直線上の点集合についても長さを考えることができるようになる. 以下では, 実数値確率変数に対しては暗にこの可測集合および測度を仮定する. 以降は簡単のため, 実数値確率変数のみについて説明する. 定義より $\omega (\in \Omega)$ ではなく $\{\omega\} (\subset \Omega)$ が事象であり, 確率変数 x の定義域は \mathcal{E} ではなく Ω であることに注意して実現値の傾向を特徴づけるのが, つぎに定義する分布である.

定義 A.5 (分布・法則). 実数値確率変数 x に対して, $\mathcal{B}(\mathbb{R})$ 上の関数 \mathbb{P}_x を

$$\mathbb{P}_x(B) := \mathbb{P}(x^{-1}(B)), B \in \mathcal{B}(\mathbb{R})$$

で定義する. このとき, $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}_x)$ は確率空間となり, 確率測度 \mathbb{P}_x を x の分布あるいは法則と呼ぶ. \triangleleft

[†] ルベーグ測度の具体的な構成方法については例ええば [] を参照のこと.

$\mathcal{B}(\mathbb{R})$ は \mathbb{R} の任意の区間を含むのであった。よって、例えば $B = [a, b]$, $a < b$ ととれば、 $x \in [a, b]$ となる確率は $\mathbb{P}_x([a, b])$ で与えられるわけである。以上、慣れ親しんだ概念・用語・記号を抽象的に再導入した。逆像を用いて様々な事象（特に確率 0 の事象）を適切に扱えることが集合論にもとづいた定式化の特徴である。

(2) ルベーグ積分 1.2 節では、確率密度関数 φ_x の存在性を仮定して期待値 $\mathbb{E}[x]$ を定義した。一方、定義 A.3 では、確率密度関数は導入していない。そこで、これらの違いを埋める役割を果たす、測度空間 $(\Omega, \mathcal{E}, \mu)$ における可測関数 f のルベーグ積分を導入する。リーマン積分は縦の短冊切り、ルベーグ積分は横の短冊切りのイメージ図を用いて説明されることが多い。これは、可測関数がある種の階段関数により近似できることを示すつぎの定理が中心的な役割を果たすためである。

定理 A.6 (単関数による可測関数の下からの近似)。指示関数 $\mathbb{1}_{E_i}$ を用いて

$$f(\omega) = \sum_{i=1}^k a_i \mathbb{1}_{E_i}(\omega), \quad a_i \in \mathbb{R}, E_i \in \mathcal{E}, \quad i = 1, 2, \dots, k \text{ は互いに素}, \quad \text{A.1}$$

と書ける可測関数 f を单関数と呼ぶ。また、任意の非負可測関数 f に対して、ある非負单関数（すなわち A.1 において $a_i \geq 0$ ）の単調増加列 $\{f_n\}_{n=1,2,\dots}$ が存在して、任意の $\omega \in \Omega$ に対して、 $f(\omega) = \lim_{n \rightarrow \infty} f_n(\omega)$ が成り立つ。 \diamond

一般の実数値可測関数 f に対して、ルベーグ積分を定義するためには、以下の 3 ステップをふむ。

• ルベーグ積分の定義

(1) まず Ω における A.1 の单関数 f の積分を

$$\int_{\Omega} f(\omega) \mu(d\omega) := \sum_{i=1}^k a_i \mu(E_i)$$

と定義する。

(2) つぎに、一般の非負可測関数 f に対して、定理 A.6 を用いて、 Ω における f の積分を

$$\int_{\Omega} f(\omega) \mu(d\omega) := \lim_{n \rightarrow \infty} \int_{\Omega} f_n(\omega) \mu(d\omega)$$

と定める。右辺の極限は $+\infty$ も値として許せば存在し、かつ近似列 $\{f_n\}_{n=1,2,\dots}$ のとり方によらない。

- (3) 最後に、非負と限らない一般の可測関数 f に対して積分を定義する。

$f^+ := \max\{f, 0\}$, $f^- := -\min\{f, 0\}$ とおけば、 f^+ と f^- は非負可測関数であり、 $f = f^+ - f^-$ と分解できる。これを用いて、 Ω における f の積分を

$$\int_{\Omega} f(\omega) \mu(d\omega) := \int_{\Omega} f^+(\omega) \mu(d\omega) - \int_{\Omega} f^-(\omega) \mu(d\omega)$$

と定義する。ただし、右辺の 2 つの積分のうち少なくとも一方は有限とする。一般の可測集合 $E \in \mathcal{E}$ 上での積分は

$$\int_E f(\omega) \mu(d\omega) := \int_{\Omega} f(\omega) \mathbb{1}_E(\omega) \mu(d\omega)$$

で定められる。

以上でルベーグ積分が定義できた。微分積分学で学ぶ区間 $[a, b]$ におけるリーマン積分 $\int_a^b f(x) dx$ が存在するとき、これは f のルベーグ測度 μ_L に関する積分 $\int_{[a,b]} f(x) \mu_L(dx)$ と一致する。ルベーグ積分が定義できる関数の範囲はリーマン積分よりもはるかに大きい。また、リーマン積分に比べて、このほかにも数学的に便利な性質をもつのがルベーグ積分の強みである。期待値は確率変数（つまり可測関数）の確率測度 \mathbb{P} に関するルベーグ積分に他ならない。

定義 A.7 (期待値)。確率空間 $(\Omega, \mathcal{E}, \mathbb{P})$ とその上の実数値確率変数 x に対して、

$$\mathbb{E}[x] := \int_{\Omega} x(\omega) \mathbb{P}(d\omega)$$

を x の期待値と呼ぶ。 \triangleleft

確率測度 \mathbb{P} だけでなく x の分布 \mathbb{P}_x も測度であるので、実数値確率変数の \mathbb{P}_x に関する積分も定義することができる。実際、 \mathbb{P}_x と変数変換を用いることで、

期待値の積分を Ω 上の積分から \mathbb{R} 上の積分に変換することで

$$\mathbb{E}[x] = \int_{\mathbb{R}} x \mathbb{P}_x(dx)$$

が成り立つ。ここまで、測度を用いて期待値を導入したが、本文中では確率密度関数なるものの存在性を仮定していた。この二つの議論を繋げるのが、以下の定理である。

定理 A.8 (ラドン-ニコディムの定理). (Ω, \mathcal{E}) を可測空間とし、その上の測度 μ と符号付き測度 v がそれぞれ $\mu(\Omega) < \infty, |v(\Omega)| < \infty$ を満たし、かつ

$$\mu(E) = 0, E \in \mathcal{E} \text{ ならば } v(E) = 0 \quad \text{A.2}$$

が成り立つならば (v は μ に関して絶対連続であるといい、 $v \ll \mu$ で表す)，

$$v(E) = \int_E \varphi(\omega) \mu(d\omega), \forall E \in \mathcal{E} \quad \text{A.3}$$

を満たす Ω 上の非負可測関数 φ が存在する。さらに、 $\tilde{\varphi}$ も A.3 を満たすならば，

$$\mu(\{\varphi(\omega) \neq \tilde{\varphi}(\omega)\}) = 0 \quad \text{A.4}$$

であり、この意味で φ は一意に定まる。 ◇

上の定理によって定められる φ を、 v の μ に関するラドン-ニコディム微分と呼び，

$$\frac{dv}{d\mu}(\omega) \quad \text{A.5}$$

と表記する。例えば、 $v \ll \mu$ に対して、カルバック・ライブラー情報量は，

$$D_{\text{KL}}(v\|\mu) := \mathbb{E}_{\mu} \left[\log \frac{dv}{d\mu} \right] := \int_{\Omega} \log \frac{dv}{d\mu}(\omega) \mu(d\omega) \quad \text{A.6}$$

と定義される。また、条件 $\mu(\Omega) < \infty$ は緩めることができ、可測空間として $(\Omega, \mathcal{E}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ をとり、 μ をルベーグ測度としても上の結果は成立する。

この定理より、分布 \mathbb{P}_x がルベーグ測度に関して絶対連続であれば，

$$\mathbb{P}_x(B) = \int_B \varphi_x(x) \mu_L(dx), \forall B \in \mathcal{B}(\mathbb{R}) \quad A.7$$

を満たす \mathbb{R} 上の非負関数 φ_x が存在し、これを \mathbb{P}_x あるいは x の**確率密度関数**と呼ぶ。これを用いれば、実数値関数 f に対して

$$\mathbb{E}[f(x)] = \int_{\mathbb{R}} f(x) \varphi_x(x) \mu_L(dx) \quad A.8$$

が成り立ち、本文中の定義であった (1.5) と矛盾せず、定理 1.2.6 も成り立つ。

以上の議論は、ベクトル値確率変数についても同様に適用できる。たとえば、 $\varphi_{(x,y)}$ が存在するとき、 $(x,y)^{-1}(\mathbb{R} \times B_y) = y^{-1}(B_y)$ に注意すると周辺化の性質

$$\int_{\mathbb{R}} \varphi_{(x,y)}(x, y) \mu_L(dx) = \varphi_y(y)$$

も導出できる。

注意 6.1.2 で言及した期待値と極限の順序交換に関する結果をまとめておく。

定理 A.9 (単調収束定理)。非負値確率変数列 $x_k, k \in \mathbb{Z}_+$ が、

$$0 \leq x_k \leq x_{k+1} \quad A.9$$

を満たすとき、概収束極限 $\bar{x}(\omega) := \lim_{k \rightarrow \infty} x_k(\omega)$ に対して、

$$\lim_{k \rightarrow \infty} \mathbb{E}[x_k] = \mathbb{E}[\bar{x}] \quad A.10$$

が成り立つ。 \diamond

定理 A.10 (ルベーグの優収束定理)。実数値確率変数列 $x_k, k \in \mathbb{Z}_+$ に対して、ある非負確率変数 y が存在して、

$$\mathbb{E}[y] < \infty \text{かつ } |x_k| \leq y, \forall k \in \mathbb{Z} \quad A.11$$

が成り立つとする。このとき、 x_k が確率変数 \bar{x} に概収束するならば、A.10 が成り立つ。 \diamond

A.2 条件付き期待値・確率

本節では、ベイズ統計に関する概念を準備する。2.3.1節で確率密度関数の存在性を仮定して紹介した条件付けの概念を、前節の枠組みで導入する。

定理 A.11 (条件付き期待値). $(\Omega, \mathcal{E}, \mathbb{P})$ を確率空間、 \mathcal{F} を \mathcal{E} の部分 σ -加法族、すなわち \mathcal{F} は σ -加法族かつ、 $\mathcal{F} \subset \mathcal{E}$ が成り立つものとする。このとき、 $\mathbb{E}[|x|] < \infty$ を満たす実数値確率変数 x に対して

$$\int_F x(\omega) \mathbb{P}(d\omega) = \int_F g(\omega) \mathbb{P}(d\omega), \forall F \in \mathcal{F} \quad \text{A.12}$$

を満たす \mathcal{F} -可測な確率変数 g が一意に存在する。 \triangleleft

証明. $v(F) := \int_F x(\omega) \mathbb{P}(d\omega)$, $F \in \mathcal{F}$ とおくと、 v は (Ω, \mathcal{F}) 上の符号付き測度となり、測度 \mathbb{P} (を \mathcal{F} に制限したもの) に関して絶対連続になる。よってラドン-ニコディムの定理から、A.12 を満たす g が一意に存在する。 \square

定義 A.12 (条件付き期待値・確率). 定理 A.11 の g を改めて $\mathbb{E}[x|\mathcal{F}]$ と書き、 \mathcal{F} の下での x の条件付き期待値と呼ぶ。また $E \in \mathcal{E}$ に対して $\mathbb{E}[\mathbb{1}_E|\mathcal{F}]$

$$\mathbb{P}(E|\mathcal{F}) := \mathbb{E}[\mathbb{1}_E|\mathcal{F}]$$

を、 \mathcal{F} の下での E の条件付き確率と呼ぶ。 \triangleleft

ここで、期待値という名前がついているが、 $\mathbb{E}[x|\mathcal{F}]$ は確率変数であることに注意されたい。また、 Ω 上の関数 y が可測となるための最小の σ -加法族を考えると、一般的な定義との関係が明確になる。

定義 A.13. (1) (X, \mathcal{X}) を可測空間とし、 x を X -値確率変数とする。このとき、

$$\sigma(x) := \{x^{-1}(B) : B \in \mathcal{X}\}$$

は σ -加法族となり、確率変数 x が生成する σ -加法族と呼ぶ。

- (2) 同じ確率空間上の実数値確率変数 x, y に対して、 $\mathbb{E}[x|\sigma(y)]$ を y の下での x の条件付き期待値と呼ぶ。また $E \in \mathcal{E}$ に対して $\mathbb{P}(E|\sigma(y)) := \mathbb{E}[\mathbb{1}_E|\sigma(y)]$ を、 y の下での E の条件付き確率と呼ぶ。

△

これを用いれば、確率変数 x, y が与えられたとき、「事象 $y = y$ による条件付け」ではなく、 y をまとめた「確率変数 y による条件付け」のもとでの x の条件付き期待値を定義することができる。この定義のもと、定理 2.3.3 の塔特性は $\mathbb{E}[x] = \mathbb{E}[\mathbb{E}[x|\sigma(y)]]$, (2.38) の不偏性は $\mathbb{E}[x|\sigma(x)] = x$ とあらわされる。また、本書では、記号を濫用して「 $\sigma(y)$ -可測な確率変数の集合」を $\sigma(y)$ と略記していたことになる。

さらに、条件付き期待値の $\sigma(y)$ -可測性を踏まえて、ある $\bar{g}: Y \rightarrow X$ が

$$\mathbb{E}[x|\sigma(y)](\omega) = \bar{g}(y(\omega)) \quad A.13$$

を満たすとしよう。このとき、A.12において $F = y^{-1}(y)$ とすれば

$$\int_{y^{-1}(y)} x(\omega) \mathbb{P}(d\omega) = \mathbb{P}(y = y) \bar{g}(y) \quad A.14$$

が成り立つため、この $\bar{g}(y)$ を $y = y$ のもとでの x の条件付き期待値 $\mathbb{E}[x|y = y]$ と呼ぶことに違和感はないだろう。こうした一連の定義は、条件付き確率密度関数を天下り的に与える本文中の定義とも、つきの意味で整合性がとれる。

定理 A.14. 確率変数 x, y が $\mathbb{E}[|x|] < +\infty$ を満たし、 (x, y) が確率密度関数 $\varphi_{(x,y)}$ をもつとする。このとき、

$$\varphi_x(x|y = y) := \begin{cases} \frac{\varphi_{(x,y)}(x, y)}{\varphi_y(y)}, & \varphi_y(y) \neq 0 \\ 0, & \varphi_y(y) = 0 \end{cases} \quad A.15$$

と定義すると、条件付き期待値 $\mathbb{E}[x|\sigma(y)]$ と

$$\bar{g}(y) := \int_{\mathbb{R}} x \cdot \varphi_x(x|y = y) \mu_L(dx) \quad A.16$$

は A.13 を満たす。△

証明. $\mathcal{F} = \sigma(y)$ のもとで $g(\omega) := \bar{g}(y(\omega))$ が A.12 を満たすことを確認すればよい。A.8 を用いると、任意の $B_y \subset \mathbb{R}$ に対して、

$$\begin{aligned} \int_{y^{-1}(B_y)} \bar{g}(y(\omega)) \mathbb{P}(d\omega) &= \int_{B_y} \bar{g}(y) \varphi_y(y) \mu_L(dy) \\ &= \int_{\mathbb{R} \times B_y} x \cdot \varphi_{(x,y)}(x, y) \mu_L(dx, dy) = \int_{(x,y)^{-1}(\mathbb{R} \times B_y)} x(\omega) \mathbb{P}(d\omega) \end{aligned}$$

が成り立つ。これは $\int_{y^{-1}(B_y)} x(\omega) \mathbb{P}(d\omega)$ に等しい。 \square

この定理により積分 A.16 が条件付き期待値を与えるという意味で、A.15 を条件付き確率密度関数と呼ぶことが正当化される。

A.3 確率過程

確率空間に時間の概念を導入する。

定義 A.15. 確率空間 $(\Omega, \mathcal{E}, \mathbb{P})$ において、 $k \leq l$ なる任意の $k, l \in \mathbb{Z}_+$ に対して

$$\mathcal{F}_k \subseteq \mathcal{F}_l \subseteq \mathcal{E} \quad \text{A.17}$$

が成り立つとき、 σ -加法族の列 $\{\mathcal{F}_k\}_{k \in \mathbb{Z}_+}$ をフィルトレーションまたは情報増大系と呼ぶ。また、 $k \in \mathbb{Z}_+$ によりパラメetrizeされた確率変数 x_k を確率過程と呼び、任意の $k \in \mathbb{Z}_+$ に対して x_k が \mathcal{F}_k -可測であるとき、 x_k は \mathcal{F} -適合であるという。 \triangleleft

この定義を用いて、 $E \in \mathcal{F}_k$ を時刻 k までに発生する事象と解釈する。例えば、 $x_{k+1} = f(x_k, v_k)$ に対して $\mathcal{F}_k := \sigma(x_0, v_{0:k})$ を考えると、 \mathcal{F} はフィルトレーションであり、 x_k は \mathcal{F} -適合である。また、 $\mathbb{Z}_+ \cup \{\infty\}$ に値をとる確率変数 τ が $(\mathcal{F}-)$ 停止時刻であるとは、 $\tau^{-1}(\{0, 1, \dots, k\}) \in \mathcal{F}_k, \forall k \in \mathbb{Z}_+$ を満たすことをいう。例えば、 \mathcal{F} -可測な x_k によりある投資方策のもとでの時刻 k における資産をあらわす。ここで、「資産が c を超えた時点で終了する」という勝ち逃げ戦略を

$$\tau := \min\{k \in \mathbb{Z}_+ : x_k \geq c\}, \tilde{x}_k := x_{\min\{\tau, k\}} \quad \text{A.18}$$

とすると、終了するかどうかは未来の情報を利用せず判断可能、つまり τ は停止時刻である。

X -値確率過程 x_i が

$$\mathbb{P}(x_{k+1} \in B | \sigma(x_{0:k})) = \mathbb{P}(x_{k+1} \in B | \sigma(x_k)), \forall B \subset \mathcal{B}, k \in \mathbb{Z}_+ \quad \text{A.19}$$

を満たすとき、マルコフ過程であるという。また、 \mathcal{F} -適合な確率過程 x_i が $\mathbb{E}[|x_k|] < +\infty$ かつ

$$\mathbb{E}[x_{k'} | \mathcal{F}_k] = x_k, \forall k' > k \in \mathbb{Z}_+ \quad \text{A.20}$$

を満たす（両辺ともに確率変数）ときマルチングール、左辺が右辺以下であるとき優マルチングールと呼ぶ。優マルチングールはある種の単調非増加性である。実数列において下に有界な単調減少列はかならず収束することはよく知られているが、確率過程に対してつぎの結果が成り立つ。

定理 A.16. 優マルチングール x_i が

$$\inf_k \mathbb{E}[\min(x_k, 0)] > -\infty \quad \text{A.21}$$

であるとき、 x は有限の期待値をもつある確率変数 x_∞ に概収束する。 \triangleleft

定理 A.17. x_i が（優）マルチングールであるとき、任意の停止時刻 τ に対して、
 $\tilde{x}_k := x_{\min\{\tau, k\}}$ は（優）マルチングールである。 \triangleleft

この定理は、平均的に損をする方策の場合、いかなる勝ち逃げ戦略を取り込んでも平均的にはやはり損をすることを示唆している。また途中で大きく資産が増加する確率は、つぎのように上から抑えられる。

定理 A.18 (マルチングール不等式). 非負実数値優マルチングール x_i は、任意の $c > 0$ に対して、

$$\mathbb{P}\left(\sup_{0 \leq k < \infty} x_k \geq c\right) \leq \frac{\mathbb{E}[x_0]}{c} \quad \text{A.22}$$

を満たす。 \triangleleft

ここで、ディンキンの公式（定理 3.1.10）より x_i の軌道に沿った V の平均的な変化量 $\mathcal{L}_x[V](k, x) := \mathbb{E}[V(k+1, x_{k+1})|x_k = x] - V(k, x)$ が任意の k, x に対して非正であることは、 $V(x_k)$ が優マルチングールであることの十分条件である。したがって、A.22 を用いて x_k がレベルセット $\text{Lv}(V; c)$ から脱出する確率を上から評価できる。さらに、 $\text{Lv}(V; c)$ にとどまり続ける標本経路は確率 1 で

$$\lim_{k \rightarrow \infty} \mathcal{L}_x[V](k, x_k) = 0 \quad \text{A.23}$$

を満たし、かつ $V(x_k)$ は収束値をもつことも知られている。例えば、注意 3.1.11 の考察より、図 3.3bにおいて軌道が一度でも区間 $[-1, -0.2](=\text{Lv}(V_-; 0.8))$ を脱出する確率は $5/8(=V_-(x_0)/0.8)$ 以下であり、 $x_k < -0.2$ を満たし続けるすべての軌道は確かに $x_k \rightarrow -1$ ($\mathcal{L}[V_-](k, x_k) = 0.1(x_k - x_k^3) \rightarrow 0$) となっている。

A.4 定常確率過程

本節では時刻変数が $k \in \mathbb{Z}_+$ でなく $k \in \mathbb{Z}$ で与えられる確率過程 $x_k, k \in \mathbb{Z}$ の代表的な性質をまとめることとする。これまでと同様に x_i と表記し、以降では断りのない限り、 x_i は \mathbb{R}^n に値をとるものとする。2 次モーメントが有界な確率過程 x_i の自己相関関数 R_x 、自己共分散関数 C_x を

$$R_x(k_1, k_2) := \mathbb{E}[x_{k_1} x_{k_2}^\top], \quad \text{A.24}$$

$$C_x(k_1, k_2) := \mathbb{E}[(x_{k_1} - \mathbb{E}[x_{k_1}])(x_{k_2} - \mathbb{E}[x_{k_2}])^\top], \quad k_1, k_2 \in \mathbb{Z} \quad \text{A.25}$$

と定義する。また、2 次モーメントが有界な確率過程 x_i と \mathbb{R}^m -値確率過程 y_i の相互相関関数 R_{xy} 、相互共分散関数 C_{xy} を

$$R_{xy}(k_1, k_2) := \mathbb{E}[x_{k_1} y_{k_2}^\top], \quad \text{A.26}$$

$$C_{xy}(k_1, k_2) := \mathbb{E}[(x_{k_1} - \mathbb{E}[x_{k_1}])(y_{k_2} - \mathbb{E}[y_{k_2}])^\top], \quad k_1, k_2 \in \mathbb{Z} \quad \text{A.27}$$

と定める。

確率過程 x_i が**強定常**であるとは、任意の $N \in \mathbb{N}, h \in \mathbb{Z}$ と任意の N 時点 $k_1, \dots, k_N \in \mathbb{Z}$ に対して、 $(x_{k_1}, \dots, x_{k_N})$ と $(x_{k_1+h}, \dots, x_{k_N+h})$ の分布が等しいこと

をいう。つまり、任意の同時分布が時間のずれに対して不变であることを表す。

これに対して、確率過程 x_{\cdot} が弱定常であるとは、ある $\mu \in \mathbb{R}^n, R_x^{\text{st}}(\cdot) : \mathbb{Z} \rightarrow \mathbb{R}^{n \times n}$ が存在して、任意の $k, h \in \mathbb{Z}$ に対して、

$$\mathbb{E}[x_k] = \mu, \mathbb{E}[x_k x_{k-h}^\top] = R_x^{\text{st}}(h) \quad \text{A.28}$$

が成り立つことをいう。すなわち、時刻 k に依らず平均値は一定で、かつ自己相関関数は時刻のずれ h のみで決まることを表す。以降、弱定常過程の自己相関関数は、時刻のずれ h を引数として $R_x(h) = \mathbb{E}[x_k x_{k-h}^\top]$ と書く。強定常かつ有界な 2 次モーメントをもつ確率過程は弱定常であるが、逆は一般に成り立たない。

定常性は、2つの確率過程の組についても拡張できる。確率過程 x_{\cdot}, y_{\cdot} が結合強定常(弱定常)であるとは、 $z_k := [x_k^\top \ y_k^\top]^\top$ で定まる確率過程 z_{\cdot} が強定常(弱定常)であることをいう。このとき、 x_{\cdot} と y_{\cdot} の相互相関関数は時刻のずれ h のみで決まるので $R_{xy}(h) = \mathbb{E}[x_k y_{k-h}^\top]$ と書くことにする。

つぎに、平均値 $\mu \in \mathbb{R}^n$ の弱定常過程 x_{\cdot} について、

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=0}^{N-1} x_k = \mu \quad \text{A.29}$$

が成り立つとき、 x_{\cdot} は平均についてエルゴード的であるという。ただし、収束の意味は概収束や 2 次平均収束など、用いられる文脈によって異なる。エルゴード性は、十分長い時間の標本経路を一つ用意して時間平均を取れば μ が推定できるという実用上非常に便利な性質である。同様に x_{\cdot} が自己相関関数についてエルゴード的であるとは、任意の $h \in \mathbb{Z}$ に対して、

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=0}^{N-1} x_k x_{k-h}^\top = R_x(h) \quad \text{A.30}$$

が成り立つことをいう。さらに結合弱定常過程 x_{\cdot}, y_{\cdot} が相互相関関数についてエルゴード的であるとは、任意の $h \in \mathbb{Z}$ に対して、

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=0}^{N-1} x_k y_{k-h}^\top = R_{xy}(h) \quad \text{A.31}$$

が成り立つことをいう。

218 13. おわりに

つぎに、弱定常過程 x_i に対して、有限の時間幅 $N \in \mathbb{N}$ で打ち切った離散時間フーリエ変換

$$X_N(\varpi) := \sum_{k=0}^{N-1} x_k e^{-j\varpi k}, \quad \varpi \in (-\pi, \pi) \quad \text{A.32}$$

を定義できる。また、確率変数 $X_N(\varpi)$ に対して、

$$S_x(\varpi) := \lim_{N \rightarrow \infty} \frac{\mathbb{E}[X_N(\varpi) X_N^\top(-\varpi)]}{N}, \quad \varpi \in (-\pi, \pi) \quad \text{A.33}$$

が収束するならば、 S_x を弱定常過程 x_i のパワースペクトル密度行列と呼び、 x_k が \mathbb{R} 値であるときは単にパワースペクトル密度と呼ばれる。パワースペクトル密度行列がもつ意味は以下の定理で明白になる。

定理 A.19 (ヴィーナ-ヒンチンの定理)。自己相関関数 R_x をもつ弱定常過程 x_i に対して、 $\sum_{h=-\infty}^{\infty} \|R_x(h)\| < \infty$ が成り立つとする[†]。このとき、 x_i のパワースペクトル密度行列 S_x が存在し、 R_x と S_x の間に以下の関係式が成立する。

$$S_x(\varpi) = \sum_{h=-\infty}^{\infty} R_x(h) e^{-j\varpi h}, \quad \varpi \in (-\pi, \pi), \quad \text{A.34}$$

$$R_x(h) = \frac{1}{2\pi} \int_{-\pi}^{\pi} S_x(\varpi) e^{j\varpi h} d\varpi, \quad h \in \mathbb{Z}. \quad \text{A.35}$$

△

ヴィーナ-ヒンチンの定理より、パワースペクトル密度行列と自己相関関数はフーリエ変換の関係にあることが分かる。とくに x_k が \mathbb{R} に値をとるとき、

$$R_x(0) = \mathbb{E}[|x_k|^2] = \frac{1}{2\pi} \int_{-\pi}^{\pi} S_x(\varpi) d\varpi, \quad \forall k \in \mathbb{Z} \quad \text{A.36}$$

である。つまり、定常過程 (x_k) の有するパワー $\mathbb{E}[|x_k|^2]$ が角周波数域 $(-\pi, \pi)$ でどのように分布するかを示すのがパワースペクトル密度 S_x であることが分かる。

[†] $\|\cdot\|$ は行列のノルムであればどちらでもよい。証明中でも同様である。

システム同定において、同定対象の動特性に関する情報を引き出すためには、システムの全てのモードを励起できる十分豊かな入力信号を印加する必要がある。その豊かさは具体的に以下で特徴づけられる[†]。

定義 A.20. u_* を \mathbb{R}^n -値の確率過程とし、

$$r_u(h) := \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=0}^{N-1} \mathbb{E}[u_k u_{k-h}^\top], \quad h \in \mathbb{Z} \quad \text{A.37}$$

が存在するものとする。このとき、 u_* が次数 $L \in \mathbb{N}$ の PE (persistently exciting, 持続的励振) 性をもつとは、以下の行列

$$\mathcal{R}_{u,L} := \begin{bmatrix} r_u(0) & r_u(1) & \cdots & r_u(L-1) \\ r_u(-1) & r_u(0) & \cdots & r_u(L-2) \\ \vdots & \vdots & \ddots & \vdots \\ r_u(1-L) & r_u(2-L) & \cdots & r_u(0) \end{bmatrix} \in \mathbb{R}^{Ln \times Ln} \quad \text{A.38}$$

が正則であることをいう。 ◇

(u_k) が弱定常ならば、次数 L の PE 性をもつ必要十分条件は

$$\begin{bmatrix} R_u(0) & R_u(1) & \cdots & R_u(L-1) \\ R_u(-1) & R_u(0) & \cdots & R_u(L-2) \\ \vdots & \vdots & \ddots & \vdots \\ R_u(1-L) & R_u(2-L) & \cdots & R_u(0) \end{bmatrix} \quad \text{A.39}$$

が正則であることである。

例えば、部分空間同定法では同定対象のシステム次数に応じて、入力信号の PE 性次数がある値以上であることが要求される。次数 L が、入力信号のもつ同定にとっての豊かさを表すことを具体例を通して確認する。

1. 定数入力 $u_k = \bar{u} \neq 0 \in \mathbb{R}$ のとき、 $r_u(h) = \bar{u}^2$ であり、 $\mathcal{R}_{u,L}$ は $L = 1$ のときのみ正則になる。よって定数入力は 1 次の PE 性をもつ。

[†] ここで述べる PE 性の定義は一例であり、文献によっては異なる定義が採用されていることもある。

2. つぎに正弦波入力 $u_k = \sin(\omega k)$, ($\omega \neq 2\pi p, p \in \mathbb{Z}$) を考える。このとき簡単な計算から, $r_u(h) = \cos(\omega h)/2$ であることが分かる。このことから, $\mathcal{R}_{u,L}$ は $L = 1, 2$ ならば正則となり, 正弦波入力は 2 次の PE 性をもつことが分かる。同様に N 個の周波数をもつ入力 $u_k = \sum_{i=1}^p a_i \sin(\omega_i k)$, ($\omega_i \neq \omega_j, i \neq j, a_i \neq 0$) は $2p$ 次の PE 性をもつ。
3. 弱定常過程 (u_k) が $\mathbb{E}[u_k] = 0$, $\mathbb{E}[u_k u_l^\top] = Q \delta_{kl}$, ($Q \in \mathbb{R}^{n \times n}, Q > 0$) を満たす場合を考える。ただし

$$\delta_{kl} = \begin{cases} 0, & k = l, \\ 1, & k \neq l \end{cases} \quad \text{A.40}$$

である。このような (u_k) は白色雑音と呼ばれる[†]。このとき, 任意の $L \in \mathbb{N}$ について, $\mathcal{R}_{u,L} = \text{diag}(Q, Q, \dots, Q)$ は正則となる。すなわち, 白色雑音は任意の次数の PE 性をもつ。これは, 白色雑音 (u_k) のパワースペクトル密度行列が一定で, 全周波数成分を含んでいることを反映した結果であると言える。

B 定理の証明

B.1 フィッシャー情報行列

ここでは, 微分と積分の交換可能性の仮定のもと, 定理 2.3.15 の証明を与える。まず, フィッシャー情報行列と対数尤度関数のヘッセ行列の関係を与える。

補題 B.1. 任意の $x \in X$ に対して,

$$\int_Y \varphi_y(y|x=x) \text{Sc}_x(x|y=y) dy = 0 \quad \text{B.41}$$

および

[†] 白色雑音は文献によって, 違う時刻間での独立性を課すなど, 異なる定義がされていることがあるため, 注意が必要である。

$$\text{FIM}_{y|x=x} = - \int_Y \varphi_y(y|x=x) \frac{\partial^2}{\partial x^2} \log l_x(x|y=y) dy \quad B.42$$

が成り立つ。 \triangleleft

証明。 尤度関数の定義と確率密度関数の規格化条件に注意すると、

$$\int_Y \frac{\partial}{\partial x} l_x(x|y=y) dy = \frac{\partial}{\partial x} \underbrace{\int_Y \varphi_y(y|x=x) dy}_{=1} = 0 \quad B.43$$

$$\int_Y \frac{\partial^2}{\partial x^2} l_x(x|y=y) dy = \frac{\partial}{\partial x} \int_Y \left(\frac{\partial}{\partial x} l_x(x|y=y) \right)^\top dy = O \quad B.44$$

が成り立つ。ここで、微分と積分の順序を交換した。したがって、B.43 より

$$\int_Y l_x(x|y=y) \frac{\partial}{\partial x} \log l_x(x|y=y) dy = 0 \quad B.45$$

であることから、B.41 がしたがう。また、B.42 の偏微分を具体的に計算し、B.44 を用いると求める等式が導かれる。 \square

このように、 $x=x$ のもとでスコア関数 $Sc_x(x|y=y)$ の（ y の観測結果に関する）平均は 0 かつ共分散行列はフィッシャー情報行列である。以下では、つきの不等式を用いて定理 2.3.15 の証明を与える。

定理 B.2 (一般化コーチー・シュワルツ不等式)。 $z, w \in \text{rv}(\mathbb{R}^n)$ に対して、 $\mathbb{E}[ww^\top]$ が正則ならば、

$$\mathbb{E}[zz^\top] \succeq \mathbb{E}[zw^\top] \mathbb{E}[ww^\top]^{-1} \mathbb{E}[wz^\top] \quad B.46$$

が成り立つ。 \triangleleft

証明。 任意の $\alpha \neq 0 \in \mathbb{R}^n$ に対して、 $\beta := -\mathbb{E}[ww^\top]^{-1} \mathbb{E}[wz^\top] \alpha$ を

$$\mathbb{E}[(z^\top \alpha + w^\top \beta)^2] \geq 0 \quad B.47$$

に代入することで示される。 \square

定理 2.3.15 の証明. まず, $\hat{x} = \theta(y)$ とおくと, B.43 と同様に,

$$\begin{aligned} & \int_Y \varphi_y(y|x=x) S_{C_x}(x|y=y)(\theta(y) - \hat{x}_c(x))^\top dy \\ &= \int_Y \varphi_y(y|x=x) \left(\frac{\partial}{\partial x} \log l_x(x|y=y) \right) \theta^\top(y) dy \quad (\because B.41) \\ &= \frac{\partial}{\partial x} \int_Y \varphi_y(y|x=x) \theta^\top(y) dy = \frac{\partial \hat{x}_c^\top(x)}{\partial x} \end{aligned} \quad B.48$$

が得られる. したがって, $\varphi_y(y|x=x)$ を掛けて y に関して積分する操作を \mathbb{E} として定理 B.2 を適用する (B.48 は $\mathbb{E}[wz^\top]$ に対応) と,

$$\text{Var}[\hat{x}|x=x] \geq \left(\frac{\partial \hat{x}_c^\top(x)}{\partial x} \right)^\top (\text{FIM}_{y|x=x})^{-1} \left(\frac{\partial \hat{x}_c^\top(x)}{\partial x} \right) \quad B.49$$

が得られる. 求める結果は $\hat{x}_c(x) = x$ とした特別な場合である. \square

定理 2.3.15 では, バイアスを小さくするとバリアンスが小さくできないことを述べた. 一般化された B.49 からは, バリアンスを小さくするために $\hat{x}_c(x)$ が定数関数に近い (勾配が 0 に近い) \hat{x} を用いると, バイアス $\hat{x}_c(x) - x$ を任意の x に対して小さくすることはできないことがわかる.

注意 B.3. ティラー展開の 2 次項までを評価すると, 微小な $\delta_x \in X$ に対して,

$$D_{\text{KL}}(\varphi_y|_{x=x+\delta_x} \| \varphi_y|_{x=x}) = \frac{1}{2} \delta_x^\top (\text{FIM}_{y|x=x}) \delta_x + o(\|\delta_x\|^2) \quad B.50$$

が成り立つ. この関係からも, $\text{FIM}_{y|x=x}$ が小さい場合は $\varphi_y|_{x=x}$ と $\varphi_y|_{x=x+\delta_x}$ の差は小さく, y の観測結果から両者を見分けることは難しいことがわかる. \triangleleft

B.2 生成モデルの敵対的学習

本節では, 生成モデルの構築をつぎのように定式化する.

問題 B.4 (生成モデル). 与えられた次元削減する確率変数 $x \in \text{rv}(\mathbb{R}^n)$, 所望の特徴量分布をあらわす確率変数 $z_d \in \text{rv}(\mathbb{R}^m)$ および $\gamma > 0$ に対して,

$$J_g(E, D) := \mathbb{E} [\|D(E(x)) - x\|^2] + \gamma D_{\text{JS}}(E(x) \| z_d) \quad B.51$$

を最小化する E と D を求めよ. \triangleleft

この問題を解くにあたり、第二項は分布として近いことを求めており、各 x に対して $E(x)$ のとるべき値が具体的に与えられているわけではなく、取り扱いが難しい。この問題を回避するために、つぎのように問題を変換しよう。

定理 B.5 (敵対的学習)。関数 $\mathcal{D} : \mathbb{R}^m \rightarrow [0, 1]$ に対して、交差エントロピー

$$J^{\text{ce}}(E, \mathcal{D}) := \mathbb{E}_x[\log \mathcal{D}(E(x))] + \mathbb{E}_{z_d}[\log(1 - \mathcal{D}(z_d))] \quad \text{B.52}$$

および

$$\mathcal{L}(E, D, \mathcal{D}) := \mathbb{E}[\|D(E(x)) - x\|^2] + \frac{\gamma}{2} J^{\text{ce}}(E, \mathcal{D}) \quad \text{B.53}$$

を定義する。このとき、

$$\arg \min_{E, D} J_g(E, D) = \arg \min_{E, D} \max_{\mathcal{D}} \mathcal{L}(E, D, \mathcal{D}) \quad \text{B.54}$$

が成り立つ。 \triangleleft

証明。 正の定数 c_r, c_d に対して、 $c_r \log y + c_d \log(1 - y)$ を最大化する $y \in (0, 1)$ は $c_r/(c_r + c_d)$ であることは、微分により容易に確かめられる。したがって、与えられた E に対して、 $E(x) \sim \varphi_r$, $z_d \sim \varphi_d$ とそれぞれ表記すると、

$$J^{\text{ce}}(E, \mathcal{D}) = \int_{\mathbb{R}^m} \left(\varphi_r(z) \log(\mathcal{D}(z)) + \varphi_d(z) \log(1 - \mathcal{D}(z)) \right) dz$$

を最大化する \mathcal{D} は

$$\mathcal{D}_E^*(z) := \frac{\varphi_r(z)}{\varphi_r(z) + \varphi_d(z)} \in [0, 1] \quad \text{B.55}$$

で与えられる。一方、 φ_r と φ_d の距離は、

$$\begin{aligned} D_{\text{JS}}(\varphi_r \| \varphi_d) &= \frac{1}{2} D_{\text{KL}}\left(\varphi_r \middle\| \frac{\varphi_r + \varphi_d}{2}\right) + \frac{1}{2} D_{\text{KL}}\left(\varphi_d \middle\| \frac{\varphi_r + \varphi_d}{2}\right) \\ &= \frac{1}{2} \left(\log 2 + \int_{\mathbb{R}^m} \varphi_r(z) \log \frac{\varphi_r(z)}{\varphi_r(z) + \varphi_d(z)} dz \right) \\ &\quad + \frac{1}{2} \left(\log 2 + \int_{\mathbb{R}^m} \varphi_d(z) \log \underbrace{\frac{\varphi_d(z)}{\varphi_r(z) + \varphi_d(z)}}_{=1-\mathcal{D}_E^*} dz \right) \\ &= \log 2 + \frac{1}{2} J^{\text{ce}}(E, \mathcal{D}_E^*) \end{aligned} \quad \text{B.56}$$

である。したがって、

$$J_g(E, D) = \mathcal{L}(E, D, \mathcal{D}_E^*) + \gamma \log 2 \quad \text{B.57}$$

であることから求める結果がしたがう。 \square

まず、与えられた E に対して B.52 の最大化は、 $\varphi_{E(x)}(z) \gg \varphi_{z_d}(z)$ ならば $\mathcal{D}(z)$ は 1 に近い値、 $\varphi_{E(x)}(z) \ll \varphi_{z_d}(z)$ ならば 0 に近い値を出力するように要請している。このような性質を満たす \mathcal{D} は入力が $E(x)$ または z_d のいずれの標本か判別する働きをするため、識別器 (discriminator) と呼ばれる。一方、与えられた \mathcal{D} に対する B.53 の (E, D に関する) 最小化は、

$$\mathcal{L}'(E, D, \mathcal{D}) := \mathbb{E}_x \left[\|D(E(x)) - x\|^2 + \frac{\gamma}{2} \log \mathcal{D}(E(x)) \right] \quad \text{B.58}$$

の最小化と等価である。これは、各データに対して復元誤差を小さくしつつ識別器にできるだけ 0 に近い値を出力させる ($E(x)$ を z_d の標本と誤判定させる) エンコーダ・デコーダの組みを求めていく。

また、 $E_p, D_p, \mathcal{D}_{p'}$ とパラメetrizeすると、与えられた p に対して $J^{ce}(E_p, \mathcal{D}_{p'})$ を最大化する p' を求める問題、与えられた p' に対して $\mathcal{L}'(E_p, D_p, \mathcal{D}_{p'})$ を最小化する p を求める問題、のどちらにも分布間の距離 $D_{JS}(E(x) \| z_d)$ が現れない。したがって、期待値を標本近似すれば、勾配法などの標準的な最適化手法が適用できる。また、これらを交互に繰り返しとくことで、所望の性質をもつ符号化器、復号化器が得られる。

B.3 方策勾配定理

ここでは、強化学習に利用した方策勾配定理の証明を与える。

定理 9.2.6 の証明. 簡単のため V^{π_μ}, Q^{π_μ} を V, Q と表記する。ベルマン方程式と連鎖律から

$$\begin{aligned}
\frac{\partial V(x)}{\partial \mu} &= \frac{\partial}{\partial \mu} \left(\ell(x, \pi_\mu(x)) + \beta \int_{\mathbb{X}} \Psi(x'|x, \pi_\mu(x)) V(x') dx' \right) \\
&= \frac{\partial \pi_\mu(x)}{\partial \mu} \frac{\partial \ell(x, u)}{\partial u} \Big|_{u=\pi_\mu(x)} + \beta \int_{\mathbb{X}} \left\{ \frac{\partial \pi_\mu(x)}{\partial \mu} \frac{\partial \Psi(x'|x, u)}{\partial u} \Big|_{u=\pi_\mu(x)} V(x') \right. \\
&\quad \left. + \Psi(x'|x, \pi_\mu(x)) \frac{\partial V(x')}{\partial \mu} \right\} dx' \\
&= \frac{\partial \pi_\mu(x)}{\partial \mu} \frac{\partial}{\partial u} \left(\ell(x, u) + \beta \int_{\mathbb{X}} \Psi(x'|x, u) V(x') dx' \right) \Big|_{u=\pi_\mu(x)} \\
&\quad + \beta \int_{\mathbb{X}} \Psi(x'|x, \pi_\mu(x)) \frac{\partial V(x')}{\partial \mu} dx' \\
&= \frac{\partial \pi_\mu(x)}{\partial \mu} \frac{\partial Q(x, u)}{\partial u} \Big|_{u=\pi_\mu(x)} + \beta \int_{\mathbb{X}} \varphi^{\pi_\mu}(x \rightarrow x', 1) \frac{\partial V(x')}{\partial \mu} dx'. \quad B.59
\end{aligned}$$

ただし、 $\partial \pi_\mu(x)/\partial \mu$ はヤコビ行列を表し、 $\varphi^{\pi_\mu}(x \rightarrow x', k) := \varphi_{x_k}^{\pi_\mu}(x'|x_0 = x)$ とした。上式は $\partial V/\partial \mu$ についての再帰式になり、入れ子に式展開を行うと

$$\begin{aligned}
\frac{\partial V(x)}{\partial \mu} &= \frac{\partial \pi_\mu(x)}{\partial \mu} \frac{\partial Q(x, u)}{\partial u} \Big|_{u=\pi_\mu(x)} \\
&\quad + \beta \int_{\mathbb{X}} \varphi^{\pi_\mu}(x \rightarrow x', 1) \frac{\partial \pi_\mu(x')}{\partial \mu} \frac{\partial Q(x', u)}{\partial u} \Big|_{u=\pi_\mu(x')} dx' \\
&\quad + \beta \int_{\mathbb{X}} \varphi^{\pi_\mu}(x \rightarrow x', 1) \beta \int_{\mathbb{X}} \varphi^{\pi_\mu}(x' \rightarrow x'', 1) \frac{\partial V(x'')}{\partial \mu} dx'' dx' \\
&= \frac{\partial \pi_\mu(x)}{\partial \mu} \frac{\partial Q(x, u)}{\partial u} \Big|_{u=\pi_\mu(x)} \\
&\quad + \beta \int_{\mathbb{X}} \varphi^{\pi_\mu}(x \rightarrow x', 1) \frac{\partial \pi_\mu(x')}{\partial \mu} \frac{\partial Q(x', u)}{\partial u} \Big|_{u=\pi_\mu(x')} dx' \\
&\quad + \beta^2 \int_{\mathbb{X}} \varphi^{\pi_\mu}(x \rightarrow x'', 2) \frac{\partial V(x'')}{\partial \mu} dx'' \\
&= \dots \\
&= \sum_{k=0}^{\infty} \beta^k \int_{\mathbb{X}} \varphi^{\pi_\mu}(x \rightarrow x', k) \frac{\partial \pi_\mu(x')}{\partial \mu} \frac{\partial Q(x', u)}{\partial u} \Big|_{u=\pi_\mu(x')} dx'
\end{aligned}$$

が成り立つ。

□

B.4 マルコフ連鎖の漸近特性

ここでは、定理 10.2.5 を証明する。準備としてつきの命題を示す。

命題 B.6. 既約な遷移確率行列 $P \in \mathbb{R}^{n \times n}$ をもつマルコフ連鎖のある状態 $i \in \llbracket n \rrbracket$ が非周期的であるとする。このとき、 P のフロベニウス根 1 を除く全ての固有値 λ は $|\lambda| < 1$ を満たす。 \triangleleft

証明. 正方行列 P と $k \in \mathbb{N}$ に対し、 P^k の (i, j) 成分 $(P^k)_{ij}$ を $p_{ij}^{(k)}$ と略記する。

確率行列 P が既約かつ、ある状態 $i^* \in \llbracket n \rrbracket$ が非周期的なとき、 P^k が正行列となるような $k = k_0 \in \mathbb{N}$ が存在する。実際、定義 10.2.1, 10.2.2 の k_1, k_2 を用いて、 $k_0 := \max_{i, j \in \llbracket n \rrbracket} \{k_1(i, i^*) + k_2(i^*) + k_1(i^*, j)\}$ とおけば P^k が正行列となる。なぜならば、任意の $i, j \in \llbracket n \rrbracket$ に対して

$$p_{ij}^{(k_0)} \geq p_{ii^*}^{(k_1(i, i^*))} p_{i^*i^*}^{(\tilde{k})} p_{i^*j}^{(k_1(i^*, j))}, \quad \tilde{k} := k_0 - k_1(i, i^*) - k_1(i^*, j) \geq k_2(i^*)$$

が成り立つことと k_1, k_2 の定義より $p_{ij}^{(k_0)} > 0$ が成り立つからである。

P^{k_0} が正確率行列であることから、そのフロベニウス根は 1 であり、ペロン・フロベニウスの定理より単純固有値である。さらに、 P^{k_0} の 1 を除く全ての固有値の絶対値は 1 未満である。以降、 P が $\hat{\lambda} \neq 1, |\hat{\lambda}| = 1$ を満たす固有値 $\hat{\lambda}$ をもつと仮定すると矛盾が生じることを示す。

P の固有値 $1, \hat{\lambda} \neq 1$ に対応する固有ベクトルを v^*, \hat{v} とおくと、 v^*, \hat{v} は一次独立である。このとき P^{k_0} は固有値 $1, \hat{\lambda}^{k_0}$ をもち、対応する固有ベクトルは v^*, \hat{v} をとれる。ここで $|\hat{\lambda}| = 1$ より $|\hat{\lambda}^{k_0}| = 1$ である。したがって、 P^{k_0} の 1 を除く全ての固有値の絶対値は 1 未満であることを思い出すと、 $\hat{\lambda}^{k_0} = 1$ である。このとき、 P^{k_0} は固有値 1 に対して、一次独立な固有ベクトル v^*, \hat{v} をもつことになり、固有値 1 が単純であることに反する。以上より P は $|\lambda| = 1$ を満たす固有値 λ を 1 以外にもちえない。 \square

注意 B.7. 証明中で既約な確率行列 P のある状態 i が非周期的であれば、 P^k が正行列となる k が存在することを示しているが、実際は以下の 3 つが等価な条件になっている。

- (1) P^k が正行列となる $k = k_0 \in \mathbb{N}$ が存在する.
- (2) P は既約かつ, 全ての状態 i が非周期的である.
- (3) P は既約かつ, ある状態 i が非周期的である.

実際, (1) が成り立つとき, 任意の $i, j \in \llbracket n \rrbracket$ と任意の $k \geq k_0$ に対して $p_{ij}^{(k)} > 0$ が成立することから, (1) \Rightarrow (2) が成り立つ. (2) \Rightarrow (3) は明らかであり, (3) \Rightarrow (1) は既に示した. \triangleleft

この命題から, それぞれの固有空間上での漸近挙動を評価することで, 定理 10.2.5 が示される.

定理 10.2.5 の証明. P が対角化可能なときのみ示す. 一般の場合も, 固有値 1 は単純であることに注意して, ジョルダン標準形で議論すればよい. このとき, 重複を許した固有値 $\lambda = 1, \lambda_2, \dots, \lambda_n$ に対応する一次独立な固有ベクトル $\{p_{\text{st}}, v_2, \dots, v_n\}$, 左固有ベクトル $\{\mathbb{1}^\top, \bar{v}_2^\top, \dots, \bar{v}_n^\top\}$ が存在する. $\bar{v}_i^\top v_i = 1, i = 2, \dots, n$ となる長さにとり, $Q := [p_{\text{st}} \ v_2 \ \dots \ v_n]$ とおくと, $Q^{-1} = [\mathbb{1} \ \bar{v}_2 \ \dots \ \bar{v}_n]^\top$ であり, $Q^{-1}PQ = \text{diag}(1, \lambda_2, \dots, \lambda_n)$ と対角化できる. よって

$$\begin{aligned}\lim_{k \rightarrow \infty} P^k &= \lim_{k \rightarrow \infty} Q \text{diag}(1, \lambda_2^k, \dots, \lambda_n^k) Q^{-1} \\ &= Q \text{diag}(1, 0, \dots, 0) Q^{-1} \\ &= p_{\text{st}} \mathbb{1}^\top = [p_{\text{st}} \ \dots \ p_{\text{st}}].\end{aligned}$$

ここで命題 B.6 より $|\lambda_i| < 1, i = 2, \dots, n$ であることを用いた. よって任意の初期分布 p_0 に対して

$$\lim_{k \rightarrow \infty} P^k p_0 = [p_{\text{st}} \ \dots \ p_{\text{st}}] p_0 = p_{\text{st}} \quad \text{B.60}$$

が成り立つ. \square

B.5 確率勾配法

定理 B.8 (Robbins-Sigmund). 非負値確率過程 $\{z_k\}, \{a_k\}, \{b_k\}, \{c_k\}$ に対して

$$\mathbb{E}[z_{k+1} | \mathcal{F}_k] \leq (1 - a_k + b_k)z_k + c_k \quad \text{B.61}$$

が成り立つとする。このとき、

$$\sum_k a_k = \infty, \sum_k b_k < \infty, \sum_k c_k < \infty \quad \text{B.62}$$

が確率 1 で成り立つならば、 z_k は 0 に概収束する。 \triangleleft

証明. B.61 の両辺を $\prod_{i=0}^{k+1} (1 + b_i)$ で割り、

$$a_k/(1 + b_{k+1}) \mapsto a_k, c_k/\prod_{i=0}^{k+1} (1 + b_i) \mapsto c_k, z_k/\prod_{i=0}^k (1 + b_i) \mapsto z_k$$

と置き直すと、一般性を失わずに $b_k = 0$ とおける。つぎに、

$$y_k := z_k + \sum_{i=0}^{k-1} a_i z_i - \sum_{i=0}^{k-1} c_i \quad \text{B.63}$$

を定義すると、 y_k は優マルチングールである。また、停止時刻

$$\tau_d := \inf\{k \geq 0 : \sum_{i=0}^{k-1} c_i > d\} \quad \text{B.64}$$

を定義すると、 $y_{\min(k, \tau_d)}$ は下に有界 ($y_{\min(k, \tau_d)} \geq -d$) な優マルチングールとなり、定理 A.16 より任意の $d > 0$ に対して概収束する。したがって、 y_k が収束値をもつ確率は $\mathbb{P}(\tau_d = \infty)$ 以上である。さらに仮定 $\sum_k c_k < \infty$ より、 $d \rightarrow \infty$ のとき $\mathbb{P}(\tau_d = \infty) \rightarrow 1$ であるため y_k が概収束することが示された。

つぎに z_k の収束性を議論するために、

$$\sum_{i=0}^{k-1} c_i - \sum_{i=0}^{k-1} a_i z_i = z_{k+1} - y_{k+1} \geq -y_{k+1} \quad \text{B.65}$$

を考えると、 $\sum_{i=0}^{k-1} a_i z_i$ は概収束し、 z_k も概収束する。最後に、 $\sum_k a_k = \infty$ より、 z_k の収束値は 0 でなければならない。 \square

定理 B.9 (確率勾配法). 確率過程 e_{\cdot} , 非負値確率変数 $a_k \in \sigma(e_{0:k-1})$ に対して $p_{k+1} = p_k - a_k y_k$, $y_k := g(p_k) + e_k$ を考える. このとき, (11.41) が (確率 1 で) 成り立ち, (11.42) および

$$\|g(p)\|^2 \leq A + B\|p - p_*\|_2^2, \forall p \quad B.66$$

$$\mathbb{E}[e_k | e_{0:k-1} = e_{0:k-1}] = 0, \mathbb{E}[\|e_k\|^2 | e_{0:k-1} = e_{0:k-1}] < K, \forall k, e_{0:k-1} \quad B.67$$

を満たす正の定数 κ, A, B, K および $p_* \in \mathbb{R}^n$ が存在するならば, p_k は p_* に概収束する. さらに, $\{a_k\}$ が確定的な数列ならば $\lim_{k \rightarrow \infty} \mathbb{E}[\|p_k - p_*\|^2] = 0$ が成り立つ. \triangleleft

証明. y_k の定義および B.61 より,

$$\begin{aligned} \|y_k\|^2 &= \|g(p_k)\|^2 + \|e_k\|^2 + 2e_k^\top g(p_k) \\ &\leq A + B\|p_k - p_*\|^2 + \|e_k\|^2 + 2e_k^\top g(p_k) \end{aligned} \quad B.68$$

であることに注意する. また, $e_{0:k-1} = e_{0:k-1}$ のもとでは, 定理 11.2.5 の証明と同様に,

$$\begin{aligned} \|p_{k+1} - p_*\|^2 &= a_k^2 \mathbb{E}[\|y_k\|^2] - 2\mathbb{E}\left[a_k(g(p_k) + e_k)^\top (p_k - p_*)\right] + \|p_k - p_*\|^2 \\ &= a_k^2(A + B\|p_k - p_*\|^2 + K) - 2a_k g(p_k)(p_k - p_*) + \|p_k - p_*\|^2 \\ &\leq (1 - 2\kappa a_k + Ba_k^2)\|p_k - p_*\|^2 + a_k^2(A + K) \end{aligned}$$

が成り立つ. ただし, $a_k, p_k \in \sigma(e_{0:k-1})$ は定数として扱えることを用いた.

後半の $\{a_k\}$ が確定的な数列の場合, 両辺の期待値をとる ($\varphi_{e_{0:k-1}}$ を掛けて $e_{0:k-1}$ に関して積分する) ことで得られる

$$b_{k+1} \leq (1 - 2\kappa a_k + Ba_k^2)b_k + a_k^2(A + K), b_k := \mathbb{E}[\|p_{k+1} - p_*\|^2] \quad B.69$$

に補題 11.2.4 を適用すると求める結果がしたがう. 前半の結果に関しては, $z_k := \|p_{k+1} - p_*\|^2$ が $\mathbb{E}[z_k | \mathcal{F}_k] \leq (1 - 2\kappa a_k + Ba_k^2)z_k + a_k^2(A + K)$ を満たすため, 定理 B.8 を適用すると求める結果がしたがう. \square

引用・参考文献

- 1) 吉川 恒夫, 井村 順一, 現代制御論 コロナ社, 2014. [Online]. Available: <https://ci.nii.ac.jp/ncid/BB16707114.bib>
- 2) 萩原朋道, ディジタル制御入門, ser. システム制御工学シリーズ コロナ社, 1999, no. 5. [Online]. Available: <https://ci.nii.ac.jp/ncid/BA3982993X.bib>
- 3) C. M. Bishop, *Pattern Recognition and Machine Learning*, ser. Information science and statistics / series editors M. Jordan ... [et al.]. Springer, 2006. [Online]. Available: <http://ci.nii.ac.jp/ncid/BA78137415>
- 4) 伊藤 海斗, 加嶋 健司, “動的システムにおけるレイイベントモデリングとその応用—安定分布によるアプローチー”, IEICE ESS Fundamentals Review, vol. 14, no. 4, pp. 269–278, Apr. 2021. [Online]. Available: https://www.jstage.jst.go.jp/article/essfr/14/4/14_269/_article/-char/ja/
- 5) J. P. Nolan, “Bibliography on stable distributions, processes and related topics,” *aaa*, no. January 2006, pp. 1–157, 2010. [Online]. Available: [http://academic2.american.edu/\\$\sim\\\$jpnolan/stable/StableBibliography.pdf%5Cnpapers3:/publication/uuid/36A45B42-DC06-4602-8BC0-920AE4EB7253](http://academic2.american.edu/$\sim\$jpnolan/stable/StableBibliography.pdf%5Cnpapers3:/publication/uuid/36A45B42-DC06-4602-8BC0-920AE4EB7253)
- 6) 萩谷 千鳳彦, 統計分布ハンドブック, 増補版 ed. 朝倉書店, 2010. [Online]. Available: <https://ci.nii.ac.jp/ncid/BB02377860.bib>
- 7) 清水 泰隆, 統計学への確率論, その先へ : ゼロからの測度論的理解と漸近理論への架け橋 内田老鶴園, 2019. [Online]. Available: <https://ci.nii.ac.jp/ncid/BB2814143X.bib>
- 8) 井村順一, システム制御のための安定論, ser. システム制御工学シリーズ コロナ社, 2000, no. 12. [Online]. Available: <https://ci.nii.ac.jp/ncid/BA49050084.bib>
- 9) M. Kishida, and A. Cetinkaya, “Risk-Aware Linear Quadratic Control Using Conditional Value-at-Risk,” *IEEE Transactions on Automatic Control*, vol. 68, no. 1, pp. 416–423, 2022.
- 10) 杉江俊治, 藤田政之, フィードバック制御入門, ser. システム制御工学シリーズ コロナ社, 1999, no. 3. [Online]. Available: <https://ci.nii.ac.jp/ncid/BA39507117.bib>
- 11) 加嶋 健司, “ディザ量子化制御系の線形解析”, システム／制御／情報, vol. 60,

- no. 5, pp. 193–198, 2016.
- 12) 大塚 敏之, 非線形最適制御入門, ser. システム制御工学シリーズ コロナ社, 2011, no. 18. [Online]. Available: <https://ci.nii.ac.jp/ncid/BB04720647>
- 13) I. R. Petersen, “Minimax LQG control,” *International Journal of Applied Mathematics and Computer Science*, vol. 16, no. 3, pp. 309–323, 2006.
- 14) C. Belta, B. Yordanov, and E. Aydin Gol, *Formal Methods for Discrete-time Dynamical Systems*, ser. Studies in systems, decision and control / series editor Janusz Kacprzyk. Springer, 2017, no. 89. [Online]. Available: <https://ci.nii.ac.jp/ncid/BB26863546.bib>
- 15) Maciejowski Jan Marian, 足立 修一, 管野 政明, モデル予測制御 : 制約のもとでの最適制御 東京電機大学出版局, 2005. [Online]. Available: <https://ci.nii.ac.jp/ncid/BA70541836.bib>
- 16) 大塚 敏之, “非線形モデル予測制御の研究動向”, システム／制御／情報, vol. 61, no. 2, pp. 42–50, 2017.
- 17) W. Li, and E. Todorov, “Iterative linearization methods for approximately optimal control and estimation of non-linear stochastic system,” *International Journal of Control*, vol. 80, no. 9, pp. 1439–1453, Sep. 2007. [Online]. Available: <http://www.tandfonline.com/doi/abs/10.1080/00207170701364913>
- 18) Y. Tassa, N. Mansard, and E. Todorov, “Control-limited differential dynamic programming,” in *2014 IEEE International Conference on Robotics and Automation (ICRA)*, vol. 53, no. 2. IEEE, May 2014, pp. 1168–1175. [Online]. Available: <http://ieeexplore.ieee.org/document/6907001/>
- 19) R. Carmona, *Lectures on BSDEs, Stochastic Control, and Stochastic Differential Games with Financial Applications*. Philadelphia, PA: Society for Industrial and Applied Mathematics, Feb. 2016. [Online]. Available: <http://pubs.siam.org/doi/book/10.1137/1.9781611974249>
- 20) 河野 佑, 加嶋 健司, “プライバシー保護と不可観測性”, システム／制御／情報, vol. 66, no. 9, pp. 340–345, 2022.
- 21) K. Lee, Y. Umezu, K. Konno, and K. Kashima, “Observability Gramian for Bayesian inference in nonlinear systems with its industrial application,” *IEEE Control Systems Letters*, vol. 7, pp. 871–876, 2023. [Online]. Available: <https://ieeexplore.ieee.org/document/9973367/>
- 22) 足立 修一, 丸田 一郎, カルマンフィルタの基礎 東京電機大学出版局, 2012. [Online]. Available: <https://ci.nii.ac.jp/ncid/BB10368424.bib>

232 引用・参考文献

- 23) 田中 利幸, “圧縮センシングの数理”, *IEICE ESS Fundamentals Review*, vol. 4, no. 1, pp. 39–47, 2010. [Online]. Available: http://www.jstage.jst.go.jp/article/essfr/4/1/4_1_39/_article/-char/ja/
- 24) K. Hayashi, M. Nagahara, and T. Tanaka, “A User’s Guide to Compressed Sensing for Communications Systems,” *IEICE Transactions on Communications*, vol. E96.B, no. 3, pp. 685–712, 2013. [Online]. Available: https://www.jstage.jst.go.jp/article/transcom/E96.B/3/E96.B_685/_article
- 25) 永原 正章, スペースモデリング : 基礎から動的システムへの応用 コロナ社, 2017. [Online]. Available: <https://ci.nii.ac.jp/ncid/BB24561863.bib>
- 26) 太田 快人, システム制御のための数学, ser. システム制御工学シリーズ コロナ社, 2000, no. 7. [Online]. Available: <https://ci.nii.ac.jp/ncid/BA48658973.bib>
- 27) G. E. Hinton, and R. R. Salakhutdinov, “Reducing the dimensionality of data with neural networks,” *Science*, vol. 313, no. 5786, pp. 504–507, July 2006. [Online]. Available: <https://www.science.org/doi/10.1126/science.1127647>
- 28) 藤崎 泰正, “不確かなシステムの解析と設計のためのランダマイズドアルゴリズム”, *システム/制御/情報*, vol. 64, no. 1, pp. 27–32, 2020.
- 29) 矢入 健久, “機械学習とシステム同定：動的システム学習研究の動向”, *計測と制御*, vol. 58, no. 3, pp. 176–181, 2019.
- 30) 井手 剛, 矢入 健久, “機械学習技術の最近の発展とシステムモデリングへの応用”, *計測と制御*, vol. 49, no. 7, pp. 463–468, 2010.
- 31) A. C. Antoulas, *Approximation of Large-Scale Dynamical Systems*. SIAM, 2005.
- 32) 薄 良彦, “非線形力学系のクープマン作用素—最近の研究から”, *システム／制御／情報*, vol. 65, no. 8, pp. 324–329, 2021.
- 33) D. Liu, Q. Wei, D. Wang, X. Yang, and H. Li, *Adaptive dynamic programming with applications in optimal control*. Springer, 2017.
- 34) 佐々木 智丈, 加嶋 健司, “制御工学者のための強化学習入門”, *計測と制御*, vol. 58, no. 3, pp. 182–188, 2019.
- 35) S. Singh, T. Jaakkola, M. L. Littman, and C. Szepesvári, “Convergence results for single-step on-policy reinforcement-learning algorithms,” *Machine learning*, vol. 38, no. 3, pp. 287–308, 2000.
- 36) C. Szepesvári, *Algorithms for reinforcement learning*, 1st ed. Morgan & Claypool Publishers, 2010, 邦邦訳= 小山田創悟 (訳者代表・編集), 前田新一・小山雅典 (監訳) :速習強化学習—基礎理論とアルゴリズム—, 共立出版, 2017. 訳:著者 web サイトで出版後も原稿の改定が続けられ, PDF が掲載されている. 本解説記事で

は2018年6月25日版を参照した。また、邦訳には訳者による付録が追加されている。

- 37) S. J. Bradtke, "Incremental dynamic programming for on-line adaptive optimal control," Ph.D. dissertation, University of Massachusetts Amherst, 1994.
- 38) R. S. Sutton, and A. G. Barto, *Reinforcement learning: An introduction*, 2nd ed. MIT Press, 2018.
- 39) R. Kamalapurkar, P. Walters, J. Rosenfeld, and W. Dixon, *Reinforcement Learning for Optimal Feedback Control: A Lyapunov-Based Approach*. Springer, 2018.
- 40) F. L. Lewis, and D. Liu, Eds., *Reinforcement Learning and Approximate Dynamic Programming for Feedback Control*. IEEE Press, 2013.
- 41) M. Wiering, and M. van Otterlo, Eds., *Reinforcement Learning: State-of-the-Art*. Springer, 2012.
- 42) 牧野, 澄谷, 白川, Eds., これからの強化学習 森北出版, 2016.
- 43) M. A. Arbib, and J. J. Bonaiuto, Eds., *From neuron to cognition via computational neuroscience*. MIT Press, 2016.
- 44) 深井朋樹, Ed., 脳の計算論 東京大学出版会, 2009.
- 45) 銀谷賢治, “脳の目的関数とは何か”, 計測と制御, vol. 47, no. 1, pp. 69–73, 2008.
- 46) 銀谷, 石井, “学習ダイナミクスの制御と脳の物質機構”, システム/制御/情報, vol. 50, no. 8, pp. 303–308, 2006.
- 47) 銀谷賢治, 計算神経科学への招待 サイエンス社, 2007.
- 48) 田中 沙織, “遅延を伴う報酬予測の脳機構の解明: 強化学習モデルに基づく fMRI データの解析”, システム／制御／情報, vol. 53, no. 4, pp. 137–142, 2009.
- 49) E. Todorov, “Efficient computation of optimal actions,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 28, pp. 11 478–11 483, 2009.
- 50) K. Dvijotham, and E. Todorov, “Inverse optimal control with linearly-solvable MDPs,” in *ICML 2010 - Proceedings, 27th International Conference on Machine Learning*, 2010, pp. 335–342.
- 51) A. Lasota, and M. C. Mackey, *Chaos, fractals, and noise : stochastic aspects of dynamics*, 2nd ed., ser. Applied mathematical sciences. Springer-Verlag, 1994, no. v. 97. [Online]. Available: <https://ci.nii.ac.jp/ncid/BA21321720.bib>
- 52) J. R. Marden, G. Arslan, and J. S. Shamma, “Cooperative control and potential games.” *IEEE transactions on systems, man, and cybernetics. Part B, Cybernetics : a publication of the IEEE Systems, Man, and Cybernetics*

234 引用・参考文献

- Society, vol. 39, no. 6, pp. 1393–407, Dec. 2009. [Online]. Available: <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=4814554>
- 53) 岩崎徹也, LMIと制御 昭晃堂, 1997. [Online]. Available: <https://ci.nii.ac.jp/ncid/BA32829341.bib>
- 54) 蛭原義雄, LMIによるシステム制御:ロバスト制御系設計のための体系的アプローチ 森北出版, 2012. [Online]. Available: <https://ci.nii.ac.jp/ncid/BB08580379.bib>
- 55) 延山英沢, 瀬部昇, システム制御のための最適化理論, ser. システム制御工学シリーズ コロナ社, 2015, no. 21. [Online]. Available: <https://ci.nii.ac.jp/ncid/BB18920247.bib>
- 56) J. B. Lasserre, “Global optimization with polynomials and the problem of moments,” *SIAM Journal on Optimization*, vol. 11, no. 3, pp. 796–817, Jan. 2001. [Online]. Available: <http://pubs.siam.org/doi/abs/10.1137/S1052623400366802>
- 57) G. C. Pflug, and A. Pichler, *Multistage stochastic optimization*, ser. Springer series in operations research and financial engineering. Springer, 2014. [Online]. Available: <https://ci.nii.ac.jp/ncid/BB17598965.bib>
- 58) S. S. Du, B. Poczos, X. Zhai, and A. Singh, “Gradient descent provably optimizes over-parameterized neural networks,” *7th International Conference on Learning Representations, ICLR 2019*, pp. 1–19, 2019.
- 59) 鈴木大慈, “再生核ヒルベルト空間の理論によるガウス過程回帰の汎化誤差解析”, システム/制御/情報, vol. 62, no. 10, pp. 396–404, 2018.
- 60) 山本裕, システムと制御の数学, ser. システム制御情報ライブラリー / システム制御情報学会編 朝倉書店, 1998, no. 16. [Online]. Available: <https://ci.nii.ac.jp/ncid/BA35195613.bib>
- 61) 瀬戸道生, 伊吹竜也, 畠中健志, 機械学習のための関数解析入門：ヒルベルト空間とカーネル法 内田老舗, 2021. [Online]. Available: <https://ci.nii.ac.jp/ncid/BC06499594.bib>
- 62) M. P. Deisenroth, D. Fox, and C. E. Rasmussen, “Gaussian processes for data-efficient learning in robotics and control,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 2, pp. 408–423, 2015.
- 63) 大住晃, 確率システム入門, ser. システム制御情報ライブラリー / システム制御情報学会編 朝倉書店, 2002, no. 24. [Online]. Available: <https://ci.nii.ac.jp/ncid/BA56031126.bib>
- 64) B. K. Øksendal, *Stochastic Differential Equations: An Introduction with Applications*, 5th ed. Springer, 1998.

演習問題の解答

演 習 問 題

演習 2.1 x の分布によらず

$$\Phi_{Ax+\bar{\mu}}(\lambda) = \mathbb{E} \left[\exp(j\lambda^\top Ax) \right] \exp(j\lambda^\top \bar{\mu}) = \Phi_x(A^\top \lambda) \exp(j\lambda^\top \bar{\mu}) \quad \text{B.70}$$

が成り立つ。特に、 Φ_x を (2.22) とすると、定理 2.2.9 が示される。また、定理 2.2.10 の証明は $\Phi_x(\lambda)\Phi_y(\lambda)$ を具体的に計算すればよい。

演習 2.2 (2.33) より、任意の y に対して $e|_{y=y} \sim \mathcal{N}(0, \hat{\Sigma})$ である。したがって、定理 2.3.7 より e は y および $\hat{x}(y) \in \sigma(y)$ と独立である。

演習 3.1 図 3.2a は、 $x_k = x_k$ のもとでは x_{k+1} は任意の値を取りうるが、 $x_{k-1:k} = x_{k-1:k}$ のもとでは x_0, v_s および $x_{k+1} \in \sigma(x_0, v_s)$ の実現値は一意に定まるため、マルコフ過程ではない。図 3.2b は定理 3.1.4 からマルコフ過程である。

演習 3.2 遷移確率密度関数は、 $x = \pm 1$ のとき $\Psi_k(x'|x) = \delta_x(x')$ 、それ以外は

$$\Psi_k(x'|x) = \frac{1}{2c} \mathbb{1}_{[f_b(x)-c, f_b(x)+c]}(x'), \quad c := 0.5(1 - |x|) \quad \text{B.71}$$

で与えられる。

演習?? ± 1 からの距離 $V(x) = 1 - |x|$ を考えると、 $\mathcal{L}[V](k, x) < 0, \forall x \in (-1, 1)$ および $\mathcal{L}[V](k, \pm 1) = 0$ である。ここで $r > 1$ とすると、 x_k は明らかに $[-1, 1] \subset \text{Lv}(V; r)$ から出ることはないため、A.23 より初期値によらずすべての標本経路は ± 1 のいづれかに収束する。

演習?? ± 1 からの距離 $V(x) = 1 - |x|$ を考えると、 $\mathcal{L}[V](k, x) < 0, \forall x \in (-1, 1)$ および $\mathcal{L}[V](k, \pm 1) = 0$ である。ここで $r > 1$ とすると、 x_k は明らかに

236 引用・参考文献

$[-1, 1] \subset \text{Lv}(V; r)$ から出ることはないとめ、A.23 より初期値によらずすべての標本経路は ± 1 のいずれかに収束する。

演習?? 改めて $\tilde{\Pi}_0 := \Pi_0$, $\tilde{\Pi}_{k+1} := r(\tilde{\Pi}_k)$, $\tilde{\gamma}_k := \gamma(\tilde{\Pi}_k)$ と表記すると、補題 6.1.11 から、

$$\tilde{V}_i(x) := \inf_{\pi_{0:i-1}} \mathbb{E} \left[\sum_{k=0}^{i-1} \beta^k \ell(x_k, u_k) + \beta^i x_i^\top \Pi_0 x_i \middle| x_0 = x \right] = x^\top \tilde{\Pi}_i x + \tilde{\gamma}_i \quad B.72$$

であり、(6.46) と比較して、 $\Pi_k \preceq \tilde{\Pi}_k$, $\gamma_k \leq \tilde{\gamma}_k$ が成り立つ。再び、 $u_k = -K(\Pi^*)x_k$ を適用して左辺を上から評価すると、

$$\tilde{V}_i(x) \leq x^\top \Pi_i x + \gamma(\Pi_i) + \beta^i x^\top ((A - BK(\Pi^*))^i)^\top \Pi_0 (A - BK(\Pi^*))^i x \quad B.73$$

が成り立つ。ここで、 $\sqrt{\beta} \rho(A - BK(\Pi^*)) < 1$ に注意して $i \rightarrow \infty$ の極限をとると、求める結果がしたがう。

演習 10.1 ((10.13) を x について和をとると $\sum_{x,u} \zeta(x,u) = 1/(1-\beta)$)。このとき、 $\sum_u \mathbb{P}^\pi(x_k = x, u_k = u) = \mathbb{P}^\pi(x_k = x)$ に注意すると

$$\begin{aligned} \sum_u \zeta(x, u) &= \sum_{k=0}^{\infty} \beta^k \mathbb{P}^\pi(x_k = x) \\ &= \mathbb{P}^\pi(x_0 = x) + \sum_{k=1}^{\infty} \beta^k \mathbb{P}^\pi(x_k = x) \\ &= \mathbb{P}(x_0 = x) + \sum_{k=1}^{\infty} \beta^k \sum_x \Psi(x|x', u') \mathbb{P}^\pi(x_{k-1} = x', u_{k-1} = u') \\ &= \mathbb{P}(x_0 = x) + \beta \sum_{x', u'} \Psi(x|x', u') \zeta(x', u') \end{aligned}$$

演習 11.1 等式制約のみ示す。

$$\begin{aligned} \mathbb{E}^\pi \left[\sum_{k=0}^{\infty} \beta^k m_r(x_k) \right] &= \mathbb{E}[m_r(x_0)] + \beta \mathbb{E}^\pi \left[\sum_{k=0}^{\infty} \beta^k m_r(x_{k+1}) \right] \\ &= c_0 + \beta \mathbb{E}^\pi \left[\sum_{k=0}^{\infty} \beta^k \int_{\mathbb{X}} m_r(x') \Psi(x'|x_k, u_k) dx' \right] \\ &= c_0 + \beta \mathbb{E}^\pi \left[\sum_{i=0}^{\infty} \beta^i c_\psi m_r(x_k, u_k) \right] \end{aligned}$$

演習?? $\text{IG}\left(\alpha + \frac{1}{2}, \beta + (y - \mu)^2\right)$

$$\frac{\beta^\alpha}{\Gamma(\alpha)} x^{-\alpha-1} \exp\left(-\frac{\beta}{x}\right) \frac{1}{\sqrt{2\pi x}} \exp\left(-\frac{(y-\mu)^2}{2x}\right) \quad \text{B.74}$$

演習 8.1 前半の結果は

$$(L^{-\top} U)^\top \Sigma_+ (L^{-\top} U) = U^\top \underbrace{L^{-1} \Sigma_+ L^{-\top}}_{=I \text{ : (8.32)}} U = I_n \quad \text{B.75}$$

$$(L^{-\top} U)^\top \Sigma_\pm (M^{-\top} V) = U^\top \underbrace{L^{-1} \Sigma_\pm M^{-\top}}_{=UDV^\top \text{ : (8.33)}} V = D \quad \text{B.76}$$

より成り立つ。また、この等式と $\Sigma_+ = \mathbb{E}[y_+ y_+]$, $\Sigma_\pm = \mathbb{E}[y_+ y_-]$ から後半の結果も成り立つ。

演習?? ラグランジュの未定乗数 $\theta \in \mathbb{R}$ を導入し、??の代わりに

$$\int_X \varphi(x) (\ell(x) - \varepsilon \log \varphi(x)) dx + \theta \left(\int_X \varphi(x) dx - 1 \right) \quad \text{B.77}$$

を考えればよい。

演習?? 任意の $\alpha \in [0, 1]$ と $v_1, v_2 \in \text{rv}(v)$ に対して、

$$\varphi_1 := \varphi_{v_1}, \varphi_2 := \varphi_{v_2}, \varphi := \varphi_v, \varphi_3 := (1 - \alpha)\varphi_1 + \alpha\varphi_2 \quad \text{B.78}$$

と表記する。このとき、 φ_3 が正値性、規格化条件、 $\varphi_3 \ll \varphi$ を満たすことは容易に確認できる。また、関数 $x \log x$, $x \geq 0$ が x に関する凸関数であることから、

$$\begin{aligned} D_{\text{KL}}(\varphi_3 \parallel \varphi) &= \int \{\varphi_3 \log \varphi_3 - \varphi_3 \log \varphi\} dv \\ &\leq \int \left\{ (1 - \alpha)\varphi_1 \log \frac{\varphi_1}{\varphi} + \alpha\varphi_2 \log \frac{\varphi_2}{\varphi} \right\} dv \\ &= (1 - \alpha)D_{\text{KL}}(\varphi_1 \parallel \varphi) + \alpha D_{\text{KL}}(\varphi_2 \parallel \varphi) \end{aligned}$$

が成り立ち、 $D_{\text{KL}}(\cdot \parallel \varphi)$ は凸汎関数である。

さらに、 $v_1, v_2 \in \text{rv}(v; h)$ のとき、

$$\begin{aligned} D_{\text{KL}}(\varphi_3 \parallel \varphi) &\leq (1 - \alpha) \int h \varphi_1 dv + \alpha \int h \varphi_2 dv \\ &= \int h \{(1 - \alpha) \varphi_1 + \alpha \varphi_2\} dv = \int h \varphi_3 dv \end{aligned}$$

が成り立つ。

演習??

??は $(\bar{s}+1)p^{\bar{s}}$ であり, ??は $(1-\varepsilon)^{\bar{s}+1}$ となる.