

論文 / 著書情報
Article / Book Information

題目(和文)	確率的セルオートマトンの変分ベイズ学習と交通流解析への応用
Title(English)	
著者(和文)	中村 文士
Author(English)	Fumito Nakamura
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第10544号, 授与年月日:2017年3月26日, 学位の種別:課程博士, 審査員:渡邊 澄夫,新田 克己,樺島 祥介,小野 功,青西 亨
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第10544号, Conferred date:2017/3/26, Degree Type:Course doctor, Examiner:,,,,,
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

確率的セルオートマトンの
変分ベイズ学習と交通流解析への応用

東京工業大学 大学院総合理工学研究科
知能システム科学専攻

中村 文士

2016年度

Abstract

交通流において生じる現象を解析するために考案されたモデルの中で、特に確率的セルオートマトンは渋滞が生じる仕組みの解明に役立つことが知られている。従来の交通流モデルは主として現象を生成する順問題として利用されてきたが、近年ではその挙動を定めているパラメータをデータから統計的に推測する逆問題にも適用されるようになった。

しかしながら、交通流の個々の構成要素が従う確率法則が多様性を持つ場合には、統計モデルが階層的な構造を持ち多くのパラメータを含むため、従来の方法では精度の良い推定を少ない演算量で行なうことは困難であった。

本論文では、確率的セルオートマトンモデルである多種粒子 TASEP と多種粒子 ZRP の統計的推測に変分ベイズ法を適用することを提案し、その学習アルゴリズムを導出する。また提案法を人工データ、実測データに適用し他手法と比較することで、多様な要素を含む交通流モデルのパラメータの精度の良い推定を少ない演算量で実現できることを明らかにする。

目次

第1章	序論	1
1.1	背景	1
1.2	研究目的	4
1.3	論文構成	5
第2章	確率的セルオートマトン	7
2.1	CA モデル	7
2.2	交通流の代表的な DCA モデルと SCA モデル	8
2.3	TASEP	9
2.4	多種粒子 TASEP	10
2.5	ZRP	11
2.6	多種粒子 ZRP	12
第3章	TASEP の変分ベイズ学習	15
3.1	TASEP の統計的表現	15
3.2	多種粒子 TASEP の統計的表現	16
3.3	多種粒子 TASEP の変分ベイズ法の学習アルゴリズム	17
3.3.1	ベイズ法	17
3.3.2	変分ベイズ法	19
3.3.3	変分ベイズ法の学習アルゴリズム	21
3.4	計算機実験	25
3.4.1	人工データへの適用	25
3.4.2	実測データへの適用	26

3.5	考察	37
3.5.1	学習データと推定結果の整合性の確認	37
3.5.2	開放系の CA における多種粒子 TASEP の推定	38
第 4 章	ZRP の変分ベイズ学習	41
4.1	ZRP の統計的表現	41
4.2	多種粒子 ZRP の統計的表現	42
4.3	多種粒子 ZRP の変分ベイズ法の学習アルゴリズム	43
4.4	計算機実験	46
4.4.1	人工データへの適用	46
4.4.2	実測データへの適用	48
4.5	考察	56
4.5.1	推定結果の確からしさ	56
4.5.2	多種粒子 TASEP と多種粒子 ZRP の比較	60
第 5 章	他の学習法との比較	63
5.1	EM 法の学習アルゴリズム	64
5.2	ギブスサンプリングの学習アルゴリズム	65
5.3	実験設定と実験結果	67
第 6 章	結論	75

目 次

2.1	交通流における CA モデルの例	8
2.2	TASEP の動作例	10
2.3	多種粒子 TASEP の動作例	10
2.4	ZRP の動作例	11
2.5	OV 関数の例 $f(d) = 0.8(\tanh(d - 2) + \tanh(2))/(1 + \tanh(2))$. . .	11
3.1	人工データにおける混合数に対する $\mathcal{F}(p_{vb})$ の値	27
3.2	渋滞形成実験で得られた計測データを離散化し得られたデータの例	30
3.3	1 から 251 タイムステップにおける混合数に対する $\mathcal{F}(p_{vb})$ の値 ($K =$ 3 が最小)	31
3.4	251 から 501 タイムステップにおける混合数に対する $\mathcal{F}(p_{vb})$ の値 ($K = 3$ が最小)	32
3.5	501 から 751 タイムステップにおける混合数に対する $\mathcal{F}(p_{vb})$ の値 ($K = 3$ が最小)	33
3.6	751 から 1001 タイムステップにおける混合数に対する $\mathcal{F}(p_{vb})$ の値 ($K = 4$ が最小)	34
3.7	1001 から 1251 タイムステップにおける混合数に対する $\mathcal{F}(p_{vb})$ の値 ($K = 2$ が最小)	35
3.8	1251 から 1501 タイムステップにおける混合数に対する $\mathcal{F}(p_{vb})$ の値 ($K = 2$ が最小)	36
3.9	100 から 200 タイムステップにおける時空図	39
4.1	人工データを生成する真の OV 関数 f_{kj}^*	47

4.2	人工データにおける混合数に対する $\mathcal{F}(p_{vb})$ の値	48
4.3	人工データに対する OV 関数の推定結果 \hat{f}_{kj}	49
4.4	1 から 251 タイムステップにおける混合数に対する $\mathcal{F}(p_{vb})$ の値 ($K = 6$ が最小)	50
4.5	1 から 251 タイムステップにおける OV 関数の推定結果 \hat{f}_{kj}	50
4.6	251 から 501 タイムステップにおける混合数に対する $\mathcal{F}(p_{vb})$ の値 ($K = 7$ が最小)	51
4.7	251 から 501 タイムステップにおける OV 関数の推定結果 \hat{f}_{kj}	51
4.8	501 から 751 タイムステップにおける混合数に対する $\mathcal{F}(p_{vb})$ の値 ($K = 6$ が最小)	52
4.9	501 から 751 タイムステップにおける OV 関数の推定結果 \hat{f}_{kj}	52
4.10	751 から 1001 タイムステップにおける混合数に対する $\mathcal{F}(p_{vb})$ の値 ($K = 5$ が最小)	53
4.11	751 から 1001 タイムステップにおける OV 関数の推定結果 \hat{f}_{kj}	53
4.12	1001 から 1251 タイムステップにおける混合数に対する $\mathcal{F}(p_{vb})$ の値 ($K = 7$ が最小)	54
4.13	1001 から 1251 タイムステップにおける OV 関数の推定結果 \hat{f}_{kj}	54
4.14	1251 から 1501 タイムステップにおける混合数に対する $\mathcal{F}(p_{vb})$ の値 ($K = 4$ が最小)	55
4.15	1251 から 1501 タイムステップにおける OV 関数の推定結果 \hat{f}_{kj}	55
4.16	人工データの更新回数を 100 回から 1000 回に変えた時の OV 関数の推定結果 \hat{f}_{kj}	58
4.17	1 から 501 タイムステップにおける OV 関数の推定結果 \hat{f}_{kj}	59
5.1	$K = 2$ における各推定手法における汎化誤差のヒストグラム	70
5.2	$K = 2$ における各推定手法の計算時間のヒストグラム	71
5.3	$K = 4$ における各推定手法における汎化誤差のヒストグラム	72
5.4	$K = 4$ における各推定手法の計算時間のヒストグラム	73

表 目 次

2.1	TASEP, 多種粒子 TASEP, ZRP, 多種粒子 ZRP の関係	13
3.1	1 から 251 タイムステップにおけるパラメータの推定結果	31
3.2	1 から 251 タイムステップにおけるクラスタの推定結果	31
3.3	251 から 501 タイムステップにおけるパラメータの推定結果	32
3.4	251 から 501 タイムステップにおけるクラスタの推定結果	32
3.5	501 から 751 タイムステップにおけるパラメータの推定結果	33
3.6	501 から 751 タイムステップにおけるクラスタの推定結果	33
3.7	751 から 1001 タイムステップにおけるパラメータの推定結果	34
3.8	751 から 1001 タイムステップにおけるクラスタの推定結果	34
3.9	1001 から 1251 タイムステップにおけるパラメータの推定結果	35
3.10	1001 から 1251 タイムステップにおけるクラスタの推定結果	35
3.11	1251 から 1501 タイムステップにおけるパラメータの推定結果	36
3.12	1251 から 1501 タイムステップにおけるクラスタの推定結果	36
4.1	人工データを生成する真の混合比 a_k^*	47
4.2	人工データに対する混合比の推定結果 \hat{a}_k	49
4.3	1 から 251 タイムステップにおける混合比の推定結果 \hat{a}_k	50
4.4	1 から 251 タイムステップにおけるクラスタの推定結果	50
4.5	251 から 501 タイムステップにおける混合比の推定結果 \hat{a}_k	51
4.6	251 から 501 タイムステップにおけるクラスタの推定結果	51
4.7	501 から 751 タイムステップにおける混合比の推定結果 \hat{a}_k	52
4.8	501 から 751 タイムステップにおけるクラスタの推定結果	52

4.9	751 から 1001 タイムステップにおける混合比の推定結果 \hat{a}_k	53
4.10	751 から 1001 タイムステップにおけるクラスタの推定結果	53
4.11	1001 から 1251 タイムステップにおける混合比の推定結果 \hat{a}_k	54
4.12	1001 から 1251 タイムステップにおけるクラスタの推定結果	54
4.13	1251 から 1501 タイムステップにおける混合比の推定結果 \hat{a}_k	55
4.14	1251 から 1501 タイムステップにおけるクラスタの推定結果	55
4.15	1 から 251 タイムステップにおける \hat{x}_{kj} の値	57
4.16	人工データの更新回数を 100 回から 1000 回に変えた時の混合比の 推定結果 \hat{a}_k	58
4.17	1 から 501 タイムステップにおける混合比の推定結果 \hat{a}_k	59
4.18	1 から 501 タイムステップにおける \hat{x}_{kj} の値	59
4.19	2 つの推定結果における各車両に割り当てられるクラスタ	60
5.1	汎化誤差の計算における真のパラメータ	67

第1章 序論

本章では，1.1 節で研究の背景について説明を行い，1.2 節で研究の目的，1.3 節で論文の構成について述べる．

1.1 背景

交通渋滞によって引き起こされている問題は深刻であり，国土交通省の発表によると，平成 24 年度における日本の渋滞損失時間 (渋滞によって余分にかかる時間) は 1 人あたり年間約 40 時間である [87]．これは貨幣換算すると約 12 兆円の損失であり，他にも渋滞によって引き起こされる走行速度の低下による二酸化炭素の排出量が増加する問題なども含んでいる．

このような交通渋滞の原因の解明やシミュレートを行い，解消するため様々な数理モデルが提案されている．代表的なものを大別すると流体モデル，追従モデル，セルオートマトン (Cellular Automaton; CA) モデルがある．また他にも結合写像格子モデルや気体分子運動論のモデルなどがある [28, 66, 86]．

流体モデルは，個々の車両の動きを考えず，車両全体を流体とみなし，流体力学における方程式を解くことで密度，流量，速度などの関係を調べるモデルである．交通流モデルに流体モデルを初めて導入したのは，Lighthill, Whitham であり，流量と車両密度の間関係を時間と空間の偏微分方程式で記述したモデルを提案した [46]．他にも Burgers 方程式を用いたモデル [52]，ナビエストークスの方程式を基にした Kerner-Konhäuser モデルなどが提案されている [37, 41, 53, 75]．

追従モデルは，個々の車両が前方の車両の影響を受け運転方法を決定するモデルであり，前の車との車間距離や相対速度に応じた連立微分方程式により記述す

るモデルである。流体モデルが車両全体を1つの方程式で記述する巨視的なモデルであったのに対し、追従モデルは、個々の方程式を連立させた微視的なモデルである。追従モデルを最初に提案したのは、Reuschel や Pipes であり、前の車との車間距離や相対速度に応じたモデルが提案された [65, 68]。その後 Newell によって最適速度関数と呼ばれる車間距離のみに依存した関数を用いたモデルが提案され [60]、坂東らによって Newell のモデルを改良した最適速度モデルが提案されている [5]。さらに、これらのモデルを改良したモデルとして、例えば Treiber らが提案した Intelligent Driver Model などがあり、他にも様々な運転者の特性を導入したモデルが提案されている [42, 73, 74]。

CA モデルは、空間をセルと呼ばれる離散格子状に分割し、離散時間ごとにセルの状態をある規則に従って更新するモデルであり、規則が決定的なものは決定的セルオートマトン (Deterministic Cellular Automaton; DCA)、確率的なものは確率的セルオートマトン (Stochastic Cellular Automaton; SCA) と呼ばれている。CA モデルは簡単な規則のみで複雑な現象を記述できるため、化学や生物学、物理学、社会学などで応用されている。交通流の分野では、セルの状態を車両の有無として車両を規則に従って走行させるモデルである [29, 49]。追従モデルが個々の車両の動きを連続的に変化させるモデルであるのに対し、CA モデルは個々の車両の動きを離散的に変化させるモデルである。CA モデルが交通流の分野で利用されるようになったのは Nagel, Schreckenberg の NS モデルの研究からであり [54]、その後、Bhamra らによる2次元 CA モデルである BML モデル [8] や Totally Asymmetric Simple Exclusion Process (TASEP) [19]、Zero Range Process (ZRP) [18]、流体モデルを超離散化したモデルなど様々なものが提案されている [61, 62]。

これらの交通流モデルを用いて取られてきた解析方法は、実測により得られる交通現象を表す指標がモデルによって生成できるように改良するというものであった。交通現象を表す指標として代表的なものに Q-K 図, Q-V 図, K-V 図などがある [75, 92]。Q は流量で、ある地点を単位時間当たり通過する車両の数を表す量である。K は密度で、ある区間上に存在する車両の数を表す量である。V は空間平均速度で、ある区間上に存在する車両の平均的な速度である。それらの関係を表

したものがQ-K図などであり、特にQ-K図は基本図と呼ばれる。実測の基本図は臨界密度を持ち、密度が臨界密度より小さい場合流量は密度に対し正比例するが、臨界密度より大きい場合流量は減少する。この性質は様々な実測データで観測されており普遍性を持つため [11, 24, 26, 40]、物理学の枠組みで基本図を用いた交通流モデルの解析がしばしば行われている [6, 15, 16, 34, 59]。

しかしながら、この解析方法は現象を生成する順問題的な方法であり、モデルの最適化を体系的に構築する方法は確立されていなかった。すなわち、モデルの有するパラメータは経験的に定められ、基本図などの交通指標の再現は定性的に行われていた。そのため、それぞれの実測データに対して再現性や精度の低いモデル構築が行われてしまうという問題があった。

この問題に対して近年、データからSCAモデルが有するパラメータを統計的に推測する逆問題的な方法が提案されている。小林、山崎はSCAモデルの一種であるStochastic Optimal Velocity(SOV)モデルのパラメータを基本図の特徴量から最尤法による方法 [90] や別のSCAモデルでZero Range Process(ZRP)と呼ばれるモデルのパラメータの最尤法と車両密度の最適化をする方法を提案している [43]。さらに、山崎は多種粒子TASEPと呼ばれるSCAモデルのベイズ法と最尤法によるクラスタリングで車両のグループ分けを行う方法を提案している [89]。これらの研究により、体系的に精度の高いモデル構築が体系的に可能となってきている。

しかしながら、この方法による交通流モデルの研究は始まったばかりであり、モデルの汎用性や推定方法の精度や演算量の観点で問題が残されている。モデルの汎用性の観点において、SOVモデルやZRPモデルは全ての車両が単一の規則に従って走行することを想定したモデルであり、運転者の運転方法や車両の大きさなどによる規則の多様性を考慮することができず、現実の交通流を扱うには不十分であると考えられる。推定方法の観点では、多様性を考慮したモデルにおいて何種類の多様性が含まれているかということは一般的に不明であり、そのような階層構造を持つモデルに最尤法を適用するとベイズ法に比べ精度が悪くなることが知られている [27, 81, 82]。また、ベイズ法は精度良く推定することが可能であるが、推定結果を得るために必要なマルコフ連鎖モンテカルロ法(MCMC)の演算量

が多くなるという問題がある [14, 22].

そのような推定方法の問題に対して, 変分ベイズ法と呼ばれる手法が提案されている [3, 13, 23]. 変分ベイズ法は, ベイズ法の推定結果に必要な事後分布の計算を近似的に行う手法であり, モデルが混合指数型分布族に含まれている場合効率的に推定を行うことが可能である [7]. また, 混合正規分布や混合ベルヌーイ分布, 隠れマルコフモデル, 縮小ランク回帰, Latent Dirichlet Allocation など様々なモデルにおいて理論解析が行われており, それらのモデルでベイズ法と近い精度で推定できることが調べられている [31–33, 55–57, 76–80, 91, 98, 102]. さらに, ベイズ法に比べ少ない演算量で推定することが可能で, 大量のデータを高速に扱う場面においても有用である [30, 47].

1.2 研究目的

本論文では, SCA モデルの一種である多種粒子 TASEP と多種粒子 ZRP の統計的推測に変分ベイズ法を適用することを提案し, その学習アルゴリズムを導出する. また, 有効性を人工データ, 実測データへの適用と他の学習法との比較により検証する.

TASEP は車両の平均速度をパラメータとして有するモデルで, ZRP は車間距離に応じて異なる前進確率をパラメータに持つモデルである. それらのモデルが多様性を持つ, つまり複数の平均速度や車間距離に対する前進確率をパラメータとして持つモデルが多種粒子モデルであり, 本論文ではそれらのパラメータの推定を行う. パラメータ推定を行うため, 学習データとして各車両が CA 上を時間発展することで得られる時空図を用いる. この時空図データを生成する確率分布を本論文では統計的表現と呼び, それぞれのモデルの統計的表現を構成し, それらの分布のパラメータ推定を行う. また, 今回考える統計的表現は混合指数型分布族に含まれており, 変分ベイズ法による効率的な推定アルゴリズムを導出することが可能であるため, そのアルゴリズムの導出を行う. 得られた提案法を人工データに適用し正しく推定が行われることの検証を行う. 検証方法として, 本論

文では人工データを生成する真のパラメータが有する多様性の種類を変分ベイズ法におけるモデル選択で選ぶことが可能であるかということ，得られたモデルで真のパラメータに近い値を推定できるかということを調べる．また，実測データに適用し多様性を持ったモデルが抽出可能であるかということを検証する．さらに，他の学習法との比較のため，予測精度を表す指標である汎化誤差と学習に要した時間を測ることで，最尤法，ベイズ法，変分ベイズ法の3つの手法について比較を行い，予測精度の良い推定を少ない演算量で実現できることを明らかにする．

本論文で用いる多種粒子 TASEP や多種粒子 ZRP は，運転方法の多様性を考慮した交通流モデルの中で基本的なモデルであるが，複雑な運転方法を表現可能なモデルの多くは，今回の推定方法と類似したアルゴリズムにより推定可能であるため，本論文は多様な運転特性を変分ベイズ法により推定するための基礎となる研究である．

1.3 論文構成

本論文の構成は以下のとおりである．2章で代表的な CA モデルと本論文で推定を行う SCA モデルについて挙動の説明を行う．3章で多種粒子 TASEP の変分ベイズ法による推定アルゴリズムの導出と人工データ，実測データへ適用した結果について述べ，4章で多種粒子 ZRP の変分ベイズ法による推定アルゴリズムの導出と人工データ，実測データへ適用した結果について述べる．5章で最尤法，ベイズ法，変分ベイズ法の3つの手法の予測精度，演算量の比較を行い，6章で結論を述べる．

第2章 確率的セルオートマトン

本章では，2.1 節で CA モデルの基礎概念について説明を行い，2.2 節で交通流解析で用いられる代表的な CA モデルを紹介し，2.3-2.6 節で本論文で用いる CA モデルの挙動について説明する．

2.1 CA モデル

セルオートマトン (CA) モデルは空間をセルと呼ばれる離散格子状に分割し，離散時間ごとに各セルの状態をある規則に従って更新するモデルである．規則によって決定的なものと確率的なものがあり，決定的なものを決定的セルオートマトン (DCA)，確率的なものを確率的セルオートマトン (SCA) と呼ぶ．CA モデルは簡単な規則のみで複雑な現象を表現することができるため，化学や生物学，物理学，社会学などに応用されている．例えば，化学の分野では Belousov-Zhabotinsky 反応 (BZ 反応) の様子を CA モデルにより表現する研究 [48, 50]，生物学の分野では貝の貝殻の模様を CA モデルにより表現する研究 [1, 45]，社会学の分野では人の動きを CA により表現する研究 [9, 83] や災害時の避難シミュレーションのための CA モデルの研究などが行われている [10, 63]．

交通流の分野では，各セルの状態が各車両の有無を表現し，車両の動きの規則を定めることで車の流れを表現する．本論文では信号機や分岐・合流・車の追い越しのない 1 車線の道路を考え，セルの端と端がつながっている周期境界条件を持ち，各車両は同一方向に同時更新することで移動する（パラレル・アップデート）場合を考える．また，1つのセルは最大1つの車両が占有できるという体積排除効果を持つものとする．つまり，ある車両に注目すると，その車両が占有する

セルの前のセルが別の車両によって占有されているとき、次の時刻で同じセルに留まる。図 2.1 にセルの全長を $L = 8$ 、全時間を $T = 4$ とした場合の CA モデルの時間発展の様子を示す。このような CA モデルの時間発展のことを時空図と呼ぶ。各区画がセルを表しており、セルの中の数字がセルの状態を示している。セルの

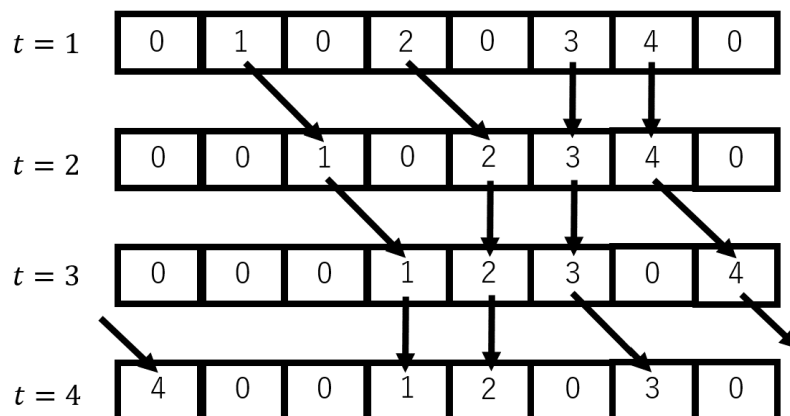


図 2.1: 交通流における CA モデルの例

値が $s = 0$ の場合、車両が存在しないことを表しており、 $s = 0$ より大きい場合、それぞれ車両が存在することを表し、同じ値が同じ車両を表現している。 $t = 2$ における車両番号 2, 3 の車両のように前のセルが別の車両によって占有されている場合、体積排除効果により前進できず同じセルに留まる。また、周期境界条件により $t = 3$ における車両番号 4 の車両は 8 番目のセルから 1 番目のセルに前進している。

2.2 交通流の代表的な DCA モデルと SCA モデル

ここでは、交通流における CA モデルとして代表的な DCA モデルと SCA モデルを紹介する。DCA モデルの代表的なモデルに Rule184 や Slow-to-Start(SIS) モデル、Fukui-Ishibashi(FI) モデルなどがある。Rule184 は決定的なものの中で最も単純なモデルであり、Wolfram が考案した 1 次元 CA モデルである Elementary

Cellular Automata の 1 つである [84, 85]. このモデルは一つ前のセルが空いている場合, 次の時刻で前進し, 空いていない場合同じセルに留まるというものである. SIS モデルは高安, 高安によって提案されたもので, Rule184 を拡張し, 1 度停止した車両は 1 タイムステップ遅延して前進するモデルである [72]. FI モデルは福井, 石橋によって提案されたもので, SIS モデルと同様に Rule184 を拡張したモデルで, 1 タイムステップで進めるセルの数が前に空いているセルの数と最大速度の最小値となるモデルである [21].

SCA モデルの代表的なモデルに Nagel・Schreckenberg(NS) モデル, Totally Asymmetric Simple Exclusion Process(TASEP) や Zero Range Process(ZRP), Stochastic Optimal Velocity(SOV) などがある. NS モデルは Nagel, Schreckenberg によって提案されたもので, 1 タイムステップで進むセルの数を加速・減速・ランダム減速によって決定するモデルである [54]. SOV は金井らによって提案されたモデルで, 追従モデルにおける OV モデルを離散化したモデルであり, 1 タイムステップ前の速度に応じて現在の速度が変化し, 速度に応じた確率で前進するモデルである [35]. TASEP, ZRP については次の節以降でそれぞれ詳しく説明する.

2.3 TASEP

TASEP は Rule184 における前進を確率的に行うモデルである. つまり, 1 つ前のセルが空いている場合, 確率 f で 1 つ前のセルに前進し, 確率 $1 - f$ で留まるモデルである. TASEP に従う CA の時間発展を図 2.2 に示す. この場合, 車両番号 1 の車両は確率 f で進んでいて, 車両番号 3 の車両は前が空いているが確率 $1 - f$ で同じセルに留まっている. この確率 f のことをホップ確率と呼び, f が大きいほど前に進みやすく, 小さいほど前に進みにくいことを表しているため, 車両の平均速度を表す値である. TASEP は SCA の中で最も単純なモデルであるため, 様々な解析が行われており, 例えば基本図と呼ばれる車両密度と流量との関係図について理論的な解析が行われている [2, 20, 36].

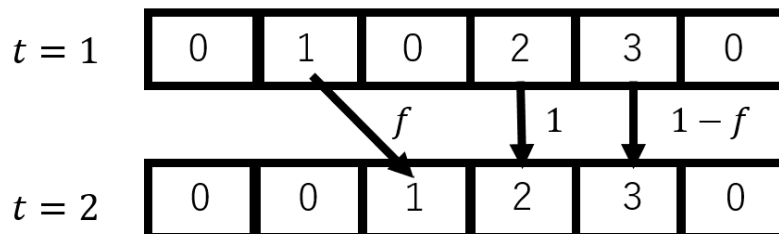


図 2.2: TASEP の動作例

2.4 多種粒子 TASEP

TASEP が全ての車両が同じホップ確率で走行するモデルであったのに対し，多種粒子 TASEP は，各車両が K 種類のホップ確率 f_1, f_2, \dots, f_K のうちどれかに従って走行するモデルである．多種粒子 TASEP に従う CA の時間発展を図 2.3 に示す．この場合，車両番号 1 と 4 の車両はホップ確率 f_1 に従い，車両番号 2 と 3 は f_2 に

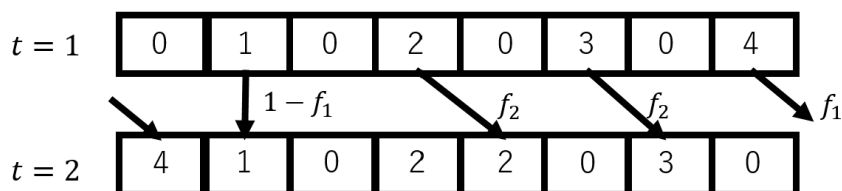


図 2.3: 多種粒子 TASEP の動作例

従っている．このような運転方法の多様性を考慮することは，バスやトラックのように重い車や乗用車のように軽い車，さらに乗用車の中でも速い車や遅い車など様々な運転方法で走行する車が内在した状況を再現できるため，現実の交通流を扱う上で重要である [4, 12, 51, 69]．特に多種粒子 TASEP に関連する研究として，例えば Karimipour は本論文で考えた多種粒子 TASEP で粒子間の追い越しがある場合について基本図の解析を行っており [38]，Rákos, Schütz は追い越しがなく，全ての車両のホップ確率が異なる多種粒子 TASEP について解析を行っている [67]．

2.5 ZRP

ZRP は車間距離 d によって前進確率 $f(d)$ が変化するようなモデルである．ここで，車間距離とは前の車両とのセルの数である．図 2.4 に ZRP の動作の例を示す．この場合，車両番号 1 の車両は車両番号 2 の車両と車間距離が 1 であるため，確率

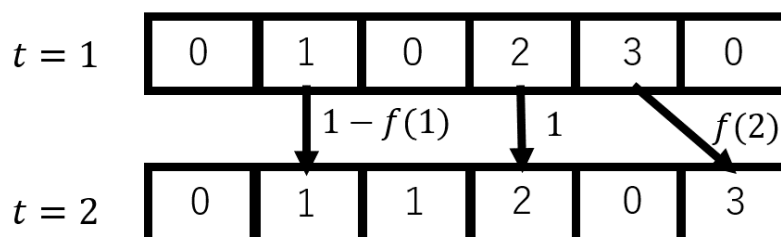


図 2.4: ZRP の動作例

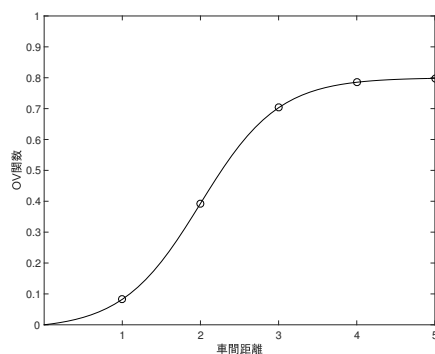


図 2.5: OV 関数の例 $f(d) = 0.8(\tanh(d - 2) + \tanh(2))/(1 + \tanh(2))$

$f(1)$ に従って前進するかどうかが決まっている．車両番号 3 の車は周期境界条件により車両番号 1 の車が前の車であり，車間距離が 2 であるため，確率 $f(2)$ で前進するかどうかが決まっている．車両番号 2 の車は車両番号 3 の車と車間距離が 0 であり，体積排除効果によって確率 1 で同じセルにとどまっている．したがって $f(0) = 0$ である．

この車間距離 d に対する前進確率の関数 $f(d)$ のことを交通流の分野では最適速度 (Optimal Velocity; OV) 関数と呼ぶ．図 2.5 のように，OV 関数によって車間の

詰めやすさや OV 関数の最大値である目標速度などを表現することができるため、TASEP に比べ表現力の高いモデルである。また、OV 関数が車間距離に対して一定である場合、TASEP と同じモデルを表すため、ZRP は TASEP を拡張したモデルである。TASEP の場合と同じく ZRP に関する解析も様々な事柄が研究されており、例えば基本図に関する理論解析などが行われている [16, 34, 39].

なお、OV 関数は単調増加性が仮定されることが多く、例えばパラメータを c とし以下のようなシグモイド型の関数などが用いられる。

$$f(d, c) = \frac{\tanh(d - c) + \tanh(c)}{1 + \tanh(c)}. \quad (2.1)$$

本論文では、推定を行いやすくするため、単調増加性を仮定せず、各車間距離で OV 関数が $[0, 1]$ の間の値を取ることが可能な場合を考える。

2.6 多種粒子 ZRP

TASEP における多種粒子 TASEP と同じように、多種粒子 ZRP は各車両が K 種類の OV 関数 $f_1(d), \dots, f_K(d)$ のうちどれか一つに従って走行するモデルである。したがって、車間の詰めやすさや目標速度に関して多様性を含むことが可能なモデルであり、前述の3つのモデルに比べより詳細な運転方法を表現可能なモデルである。ここで、各 OV 関数が車間距離に対して一定である場合、多種粒子 TASEP と同じモデルを表すため、多種粒子 ZRP は多種粒子 TASEP を拡張したモデルである。多種粒子 ZRP に関する研究として Evans, Hanney などにより解析が行われている [17, 25].

ここまでの4つのモデルの関係をまとめると、表 2.1 のようになる。3章で表の左側についてパラメータを統計的推測する方法について議論を行い、4章で表の右側について検討する。

表 2.1: TASEP, 多種粒子 TASEP, ZRP, 多種粒子 ZRP の関係

	車間距離を考慮しない	車間距離を考慮する
多様性を含まない	TASEP	ZRP
多様性を含む	多種粒子 TASEP	多種粒子 ZRP

第3章 TASEP の変分ベイズ学習

パラメータ推定を行う多くのアルゴリズムでは、データを生成する確率モデルを定義し、それを基にパラメータの最適化を行う。本論文では、時空図データを生成する確率モデルを統計的表現と呼ぶ。本章では、まず3.1節で全ての車両が同じホップ確率に従う(通常の)TASEPの統計的表現を紹介し、3.2節で複数のホップ確率が混在する多種粒子TASEPの統計的表現を述べる。その後、3.3節で変分ベイズ法の学習アルゴリズムがどのようなものであるか説明を行い、多種粒子TASEPに適用した時に得られる学習アルゴリズムを導出する。3.4節で得られた学習アルゴリズムを人工データ、実測データに適用した結果を示し、3.5節で考察を行う [58].

3.1 TASEP の統計的表現

TASEPではホップ確率 f がパラメータである。ある時刻で前のセルが空いている場合、確率 f で「前進する」という事象が起き、確率 $1 - f$ で「同じセルに残る」という事象が起こる。これらの事象は前の時刻の事象に依らず独立に起こるため、ベルヌーイ試行である。「前に進む」という事象を「成功」、「同じセルに残る」という事象を「失敗」とみなすと、ある車両の試行回数(「成功」と「失敗」の合計回数) x が与えられたもとで「成功」の回数 y の分布は、ホップ確率 f を用いて2項分布で表すことができる。

$$p(y|x) = \binom{x}{y} f^y (1-f)^{x-y}. \quad (3.1)$$

ここで、 $\binom{x}{y}$ は2項係数である。これが全ての車両が同じホップ確率に従うTASEPの統計的表現である。この x と y は、CAの時間発展である時空図から数えること

ができ、 i 番目の車両の試行回数を $x^{(i)}$ 、成功回数を $y^{(i)}$ とする。また、 n 台の車両の試行回数と成功回数の集合をそれぞれ $x^n = \{x^{(1)}, \dots, x^{(n)}\}$ 、 $y^n = \{y^{(1)}, \dots, y^{(n)}\}$ で表す。この時、式 (3.1) をホップ確率も含めて $p(y|x, f)$ で表すと、TASEP の統計的表現による尤度 $p(y^n|x^n, f)$ は

$$p(y^n|x^n, f) = \prod_{i=1}^n \binom{x^{(i)}}{y^{(i)}} f^{y^{(i)}} (1-f)^{x^{(i)}-y^{(i)}}$$

となる。したがって、最尤推定量 \hat{f} は経験確率に一致し、

$$\hat{f} = \arg \max_f p(y^n|x^n, f) = \frac{\sum_{i=1}^n y^{(i)}}{\sum_{i=1}^n x^{(i)}}$$

となる。

3.2 多種粒子 TASEP の統計的表現

多種粒子 TASEP では、 K 種類のホップ確率 f_1, \dots, f_K が比率 a_1, \dots, a_K の割合で割り当てられる。割り当てられるホップ確率の番号を表現するため、隠れ変数 $Z = (Z_1, \dots, Z_K)$ を導入する。この隠れ変数は以下のような集合に値をとる確率変数である。

$$Z \in \mathcal{C} = \{(1, 0, \dots, 0), (0, 1, \dots, 0), (0, 0, \dots, 1)\}.$$

この集合の要素は Z_k の中のどれか一つが 1 となり、その他が 0 となる。この隠れ変数を用いて、割り当てられるホップ確率の番号の確率は比率 (a_1, \dots, a_K) を用いて

$$p(z) = \prod_{k=1}^K a_k^{z_k} \quad (3.2)$$

で表すことができる (ただし $\sum_{k=1}^K a_k = 1$)。また、隠れ変数 z と試行回数 x が与えられている場合、成功回数が y である確率は

$$p(y|x, z) = \prod_{k=1}^K \left\{ \binom{x}{y} f_k^y (1-f_k)^{x-y} \right\}^{z_k} \quad (3.3)$$

となる．式 (3.2),(3.3) を用いて成功回数と隠れ変数の同時分布 $p(y, z|x)$ は

$$p(y, z|x) = p(y|x, z)p(z) = \prod_{k=1}^K \left\{ a_k \binom{x}{y} f_k^y (1 - f_k)^{x-y} \right\}^{z_k} \quad (3.4)$$

である．隠れ変数は観測できないため，全ての可能性を考慮し，式 (3.4) を隠れ変数について周辺化することで多種粒子 TASEP の統計的表現が得られ，以下の混合分布となる．

$$p(y|x) = \sum_{z \in \mathcal{C}} p(y, z|x) = \sum_{k=1}^K a_k \binom{x}{y} f_k^y (1 - f_k)^{x-y}. \quad (3.5)$$

多種粒子 TASEP において，パラメータは各車両に割り当てられるホップ確率の比率 $a_k (k = 1, \dots, K)$ とホップ確率の値 $f_k (k = 1, \dots, K)$ である．これらをまとめて $w = (a_k, f_k; k = 1, \dots, K)$ で表す．本章では，以後パラメータを明記する場合，式 (3.5) は $p(y|x, w)$ とする．なお，全ての $x^{(1)} = x^{(2)} = \dots = x^{(n)}$ であるとき，これは混合 2 項分布に一致する．それゆえ，多種粒子 TASEP は混合 2 項分布を拡張したものである．

3.3 多種粒子 TASEP の変分ベイズ法の学習アルゴリズム

多種粒子 TASEP のパラメータ推定を行うため，ここでは変分ベイズ法と呼ばれる方法を用いる．変分ベイズ法は，ベイズ法の近似アルゴリズムであるため，まずベイズ法について説明を行い，その後，変分ベイズ法の説明と学習アルゴリズムの導出を行う．

3.3.1 ベイズ法

ベイズ法は，多種粒子 TASEP モデルの隠れ変数との同時分布である式 (3.4) と n 個のデータ (x^n, y^n) ，パラメータ w の事前分布 $\varphi(w)$ を用いる．ここでは事前分

布として

$$\varphi(w) = \text{Dir}(a|\{\phi_k\}_{k=1}^K) \prod_{k=1}^K \text{Beta}(f_k|\alpha_k, \beta_k) \quad (3.6)$$

を用いる. ここで, $\text{Dir}(\cdot)$, $\text{Beta}(\cdot)$ はそれぞれディリクレ分布, ベータ分布を表しており, それぞれガンマ関数 $\Gamma(\cdot)$ を用いて

$$\begin{aligned} \text{Dir}(a|\{\phi_k\}_{k=1}^K) &= \frac{\Gamma(\sum_{k=1}^K \phi_k)}{\prod_{k=1}^K \Gamma(\phi_k)} \prod_{k=1}^K a_k^{\phi_k-1}, \\ \text{Beta}(f_k|\alpha_k, \beta_k) &= \frac{\Gamma(\alpha_k + \beta_k)}{\Gamma(\alpha_k)\Gamma(\beta_k)} f_k^{\alpha_k-1} (1-f_k)^{\beta_k-1} \end{aligned}$$

である. また, $\phi_k, \alpha_k, \beta_k > 0$ はハイパーパラメータであり, 本論文では $\phi_1 = \dots = \phi_K = \phi, \alpha_1 = \dots = \alpha_K = \alpha, \beta_1 = \dots = \beta_K = \beta$ とする. 各車両に割り当てられるホップ確率の番号である n 個の隠れ変数の集合を $z^n = \{z^{(1)}, \dots, z^{(n)}\}$ とすると, 隠れ変数とデータの同時分布の尤度 $p(y^n, z^n|x^n, w)$ は

$$p(y^n, z^n|x^n, w) = \prod_{i=1}^n p(y^{(i)}, z^{(i)}|x^{(i)}, w) \quad (3.7)$$

によって与えられる. 尤度と事前分布を用いることで, データが与えられた時の隠れ変数とパラメータの事後分布 $p(w, z^n|x^n, y^n)$ は以下ようになる.

$$p(w, z^n|x^n, y^n) = \frac{1}{Z_n} p(y^n, z^n|x^n, w) \varphi(w). \quad (3.8)$$

ここで, Z_n は周辺尤度であり,

$$Z_n = \sum_{z^n} \int p(y^n, z^n|x^n, w) \varphi(w) dw = \int p(y^n|x^n, w) \varphi(w) dw$$

となる. この事後分布を用いて, ベイズ法は学習結果である予測分布 $p_{\text{bayes}}(y|x, x^n, y^n)$ を

$$p_{\text{bayes}}(y|x, x^n, y^n) = \sum_{z^n} \int p(w, z^n|x^n, y^n) p(y|x, w) dw \quad (3.9)$$

とする方法である. ここで, 周辺尤度 Z_n はパラメータに関する積分を含んでおり, 多種粒子 TASEP や後述の多種粒子 ZRP では解析的に計算することが困難である.

この問題に対して、マルコフ連鎖モンテカルロ法 (MCMC) を用いて事後分布からのサンプリングを行い予測分布の近似を行う方法、事後分布そのものを近似する方法としてラプラス近似、変分ベイズ法などがある。変分ベイズ法は、ラプラス近似に比べ高精度で MCMC に比べ演算量が少ない方法であるため、本論文では、変分ベイズ法を用いた方法について考える。

3.3.2 変分ベイズ法

変分ベイズ法は、今回のモデルの場合パラメータと隠れ変数が独立な近似分布 $p_{vb}(w, z^n) = p_{vb}(w)p_{vb}(z^n)$ の中で事後分布である式 (3.8) とのカルバック・ライブラ情報量 (KL 情報量) が最小になるものを選ぶ手法である。ここで、KL 情報量 $KL(p_{vb}(w, z^n)||p(w, z^n|x^n, y^n))$ は

$$KL(p_{vb}(w, z^n)||p(w, z^n|x^n, y^n)) = \sum_{z^n} \int p_{vb}(w)p_{vb}(z^n) \log \frac{p_{vb}(w, z^n)}{p(w, z^n|x^n, y^n)} dw$$

によって与えられる。この KL 情報量の $p_{vb}(w, z^n)$ への最小化は周辺尤度に依らず行うことができるため、ここで考える最小化は

$$\begin{aligned} p_{vb}(w)p_{vb}(z^n) &= \arg \min_{\bar{p}_{vb}(w)\bar{p}_{vb}(z^n)} KL(\bar{p}_{vb}(w)\bar{p}_{vb}(z^n)||p(w, z^n|x^n, y^n)) \\ &= \arg \min_{\bar{p}_{vb}(w)\bar{p}_{vb}(z^n)} \sum_{z^n} \int \bar{p}_{vb}(w)\bar{p}_{vb}(z^n) \log \frac{\bar{p}_{vb}(w)\bar{p}_{vb}(z^n)}{p(y^n, z^n|x^n, w)\varphi(w)} dw \end{aligned}$$

である。以後最後の項の汎関数を

$$\mathcal{F}(\bar{p}_{vb}) = \sum_{z^n} \int \bar{p}_{vb}(w)\bar{p}_{vb}(z^n) \log \frac{\bar{p}_{vb}(w)\bar{p}_{vb}(z^n)}{p(y^n, z^n|x^n, w)\varphi(w)} dw \quad (3.10)$$

で表記する。この汎関数を最小にする分布について以下の定理が知られている [70, 101].

定理 1. 汎関数 $\mathcal{F}(\bar{p}_{vb})$ を最小にする分布は以下の形を満たす。

$$p_{vb}(w) = \frac{1}{C_w} \exp\left(E_z[\log p(y^n, z^n|x^n, w) + \log \varphi(w)]\right), \quad (3.11)$$

$$p_{vb}(z^n) = \frac{1}{C_z} \exp\left(E_w[\log p(y^n, z^n|x^n, w)]\right). \quad (3.12)$$

ここで, C_w, C_z はそれぞれの分布の正規化定数を表し, $E_z[\cdot], E_w[\cdot]$ はそれぞれ $p_{vb}(z^n)$, $p_{vb}(w)$ の平均である. また, 式 (3.11), (3.12) の繰り返しにより $\mathcal{F}(\bar{p}_{vb})$ は単調非増加する.

証明. まず, $\bar{p}_{vb}(z^n)$ を固定すると $p_{vb}(w)$ が式 (3.11) を満たすことを示す. $\mathcal{F}(\bar{p}_{vb})$ は以下のように表すことができる.

$$\begin{aligned}\mathcal{F}(\bar{p}_{vb}) &= \sum_{z^n} \int \bar{p}_{vb}(w) \bar{p}_{vb}(z^n) \log \frac{\bar{p}_{vb}(w) \bar{p}_{vb}(z^n)}{p(y^n, z^n | x^n, w) \varphi(w)} dw \\ &= \int \bar{p}_{vb}(w) \log \bar{p}_{vb}(w) dw \\ &\quad - \int \bar{p}_{vb}(w) E_{\bar{z}}[\log(p(y^n, z^n | x^n, w) \varphi(w))] dw + \text{Const} \\ &= \int \bar{p}_{vb}(w) \log \bar{p}_{vb}(w) dw \\ &\quad - \int \bar{p}_{vb}(w) \log \exp\left(E_{\bar{z}}[\log(p(y^n, z^n | x^n, w) \varphi(w))]\right) dw + \text{Const}.\end{aligned}$$

ここで, $E_{\bar{z}}[\cdot]$ は $\bar{p}_{vb}(z^n)$ による平均であり, Const は $\bar{p}_{vb}(w)$ に依存しない項である. $\exp\left(E_{\bar{z}}[\log(p(y^n, z^n | x^n, w) \varphi(w))]\right)$ の項を確率分布にするため, 正規化定数 $C_w = \int \exp\left(E_{\bar{z}}[\log(p(y^n, z^n | x^n, w) \varphi(w))]\right) dw$ を用いると

$$\begin{aligned}\mathcal{F}(\bar{p}_{vb}) &= \int \bar{p}_{vb}(w) \log \bar{p}_{vb}(w) dw \\ &\quad - \int \bar{p}_{vb}(w) \log \frac{C_w}{C_w} \exp\left(E_{\bar{z}}[\log(p(y^n, z^n | x^n, w) \varphi(w))]\right) dw + \text{Const} \\ &= \int \bar{p}_{vb}(w) \log \bar{p}_{vb}(w) dw \\ &\quad - \int \bar{p}_{vb}(w) \log \frac{1}{C_w} \exp\left(E_{\bar{z}}[\log(p(y^n, z^n | x^n, w) \varphi(w))]\right) dw - \log C_w + \text{Const} \\ &= \int \bar{p}_{vb}(w) \log \frac{\bar{p}_{vb}(w)}{\frac{1}{C_w} \exp\left(E_{\bar{z}}[\log(p(y^n, z^n | x^n, w) \varphi(w))]\right)} dw - \log C_w + \text{Const} \\ &= KL\left(\bar{p}_{vb}(w) \parallel \frac{1}{C_w} \exp\left(E_{\bar{z}}[\log(p(y^n, z^n | x^n, w) \varphi(w))]\right)\right) - \log C_w + \text{Const}.\end{aligned}$$

したがって, $\bar{p}_{vb}(w)$ についての最小化は1項目のKL情報量の最小化によって得ら

れる. この KL 情報量の最小値は $\bar{p}_{vb}(w) = \frac{1}{C_w} \exp\left(E_{\bar{z}}[\log(p(y^n, z^n|x^n, w)\varphi(w))]\right)$ の場合に限られ, それは式 (3.11) の形から示すことができた.

次に, $\bar{p}_{vb}(w)$ を固定し, $C_z = \sum_{z^n} \exp\left(E_{\bar{w}}[\log p(y^n, z^n|x^n, w)]\right)$ を用いると, 同様の議論を $\bar{p}_{vb}(z^n)$ についても行うことができ, それは式 (3.12) の形となっているため, $\mathcal{F}(\bar{p}_{vb})$ の最小値は式 (3.11), (3.12) によって与えられることを示すことができた. また, この証明から式 (3.11), (3.12) の繰り返しにより $\mathcal{F}(\bar{p}_{vb})$ が単調非増加することも示された. \square

この定理により, それぞれの式は一方の分布の平均操作を行っており, 変分ベイズ法は一方の分布の初期値を決め, 式 (3.11), (3.12) の形を満足するように分布を交互に更新する手法である. なお, この条件は必要条件であり十分条件でないため, 一般的に式 (3.11), (3.12) を満たす分布は複数存在し, いわゆる局所解の問題を含んでいる [95]. そのため, 初期値を繰り返し変え式 (3.10) の汎関数の最小化を行う必要がある.

3.3.3 変分ベイズ法の学習アルゴリズム

ここでは, 式 (3.11), (3.12) をもとに多種粒子 TASEP に対する変分ベイズ法の学習アルゴリズムの導出を行う. 本論文では, 分布の初期値として $p_{vb}(z^n)$ の初期化を行う. $p_{vb}(z^n)$ の初期化により $E_z[z_k^{(i)}] = p(z_k^{(i)} = 1) = \hat{z}_k^{(i)}$ が全ての $i = 1, \dots, n, k = 1, \dots, K$ について得られる. 隠れ変数とデータの同時分布の対数尤度 $\log p(y^n, z^n|x^n, w)$ は

$$\begin{aligned} \log p(y^n, z^n|x^n, w) &= \sum_{k=1}^K \sum_{i=1}^n z_k^{(i)} \left\{ \log a_k + y^{(i)} \log f_k \right. \\ &\quad \left. + (x^{(i)} - y^{(i)}) \log(1 - f_k) + \log \binom{x^{(i)}}{y^{(i)}} \right\} \end{aligned}$$

で表すことができることに注意すると、 $\hat{z}_k^{(i)}$ と式 (3.6) を用いて $p_{vb}(w)$ は

$$p_{vb}(w) = \text{Dir}(a|\{\hat{\phi}_k\}_{k=1}^K) \prod_{k=1}^K \text{Beta}(f_k|\hat{\alpha}_k, \hat{\beta}_k), \quad (3.13)$$

$$\hat{\phi}_k = \hat{z}_k^{(i)} + \phi, \quad (3.14)$$

$$\hat{\alpha}_k = \sum_{i=1}^n y^{(i)} \hat{z}_k^{(i)} + \alpha, \quad (3.15)$$

$$\hat{\beta}_k = \sum_{i=1}^n (x^{(i)} - y^{(i)}) \hat{z}_k^{(i)} + \beta \quad (3.16)$$

である．次に、ディガンマ関数を $\psi(x) = \frac{d}{dx} \log \Gamma(x)$ で表記すると、パラメータの対数期待値は

$$E_w[\log a_k] = \psi(\hat{\phi}_k) - \psi\left(\sum_{k=1}^K \hat{\phi}_k\right), \quad (3.17)$$

$$E_w[\log f_k] = \psi(\hat{\alpha}_k) - \psi(\hat{\alpha}_k + \hat{\beta}_k), \quad (3.18)$$

$$E_w[\log(1 - f_k)] = \psi(\hat{\beta}_k) - \psi(\hat{\alpha}_k + \hat{\beta}_k) \quad (3.19)$$

となる．したがって、ここで得られた $(\hat{\phi}_k, \hat{\alpha}_k, \hat{\beta}_k)$ を用いると、 $p_{vb}(z^n)$ を計算することができ以下のようなになる．

$$\begin{aligned} p_{vb}(z^n) &= \frac{1}{C_z} \prod_{i=1}^n \prod_{k=1}^K \exp(z_k^{(i)} L_{ik}), \\ L_{ik} &= \psi(\hat{\phi}_k) + y^{(i)} \psi(\hat{\alpha}_k) + (x^{(i)} - y^{(i)}) \psi(\hat{\beta}_k) - x^{(i)} \psi(\hat{\alpha}_k + \hat{\beta}_k), \\ C_z &= \prod_{i=1}^n \sum_{k=1}^K \exp(L_{ik}). \end{aligned} \quad (3.20)$$

$p_{vb}(z^n)$ を用いることで再び $\hat{z}_k^{(i)}$ を計算することができ、

$$\hat{z}_k^{(i)} = E_z[z_k^{(i)}] = p(z_k^{(i)} = 1) = \frac{\exp(L_{ik})}{\sum_{l=1}^K \exp(L_{il})} \quad (3.21)$$

となる．変分ベイズ法はこの計算の繰り返しにより汎関数 $\mathcal{F}(p_{vb})$ を最小にする分布を求めるアルゴリズムである．以上をまとめると、アルゴリズムは次のようになる．

アルゴリズム 1 (多種粒子 TASEP の変分ベイズ法のアルゴリズム).

1. $\{\hat{z}_k^{(i)}; i = 1, \dots, n, k = 1, \dots, K\}$ を初期化する.

2. M-ステップ: $p_{vb}(w)$ の更新

$$\begin{aligned}\hat{\phi}_k &\leftarrow \sum_{i=1}^n \hat{z}_k^{(i)} + \phi, \\ \hat{\alpha}_k &\leftarrow \sum_{i=1}^n y^{(i)} \hat{z}_k^{(i)} + \alpha, \\ \hat{\beta}_k &\leftarrow \sum_{i=1}^n (x^{(i)} - y^{(i)}) \hat{z}_k^{(i)} + \beta.\end{aligned}$$

3. E-ステップ: $p_{vb}(z^n)$ の更新

$$\begin{aligned}L_{ik} &\leftarrow -\psi(\hat{\phi}_k) + y^{(i)}\psi(\hat{\alpha}_k) + (x^{(i)} - y^{(i)})\psi(\hat{\beta}_k) - x^{(i)}\psi(\hat{\alpha}_k + \hat{\beta}_k), \\ \hat{z}_k^{(i)} &\leftarrow \frac{\exp(L_{ik})}{\sum_{l=1}^K \exp(L_{il})}.\end{aligned}$$

4. M-ステップと E-ステップを交互に繰り返し更新する.

ここで得られたパラメータの分布 $p_{vb}(w)$ を用いることでパラメータの平均値 $\hat{w} = E_w[w] = (\hat{a}_k, \hat{f}_k; k = 1, \dots, K)$ は

$$\hat{a}_k = \frac{\hat{\phi}_k}{\sum_{l=1}^K \hat{\phi}_l}, \hat{f}_k = \frac{\hat{\alpha}_k}{\hat{\alpha}_k + \hat{\beta}_k} \quad (3.22)$$

となる. また, 予測分布 $p_{vb}(y|x, x^n, y^n)$ は

$$\begin{aligned}p_{vb}(y|x, x^n, y^n) &= \int p_{vb}(w)p(y|x, w)dw = \sum_{k=1}^K E_w[a_k] \binom{x}{y} E_w[f_k^y (1 - f_k)^{x-y}] \\ &= \sum_{k=1}^K \hat{a}_k \binom{x}{y} \frac{\text{Beta}(\hat{\alpha}_k + y, \hat{\beta}_k + (x - y))}{\text{Beta}(\hat{\alpha}_k, \hat{\beta}_k)}\end{aligned}$$

である. ここで $\text{Beta}(\cdot, \cdot)$ はベータ関数を表す.

さらに, 汎関数 $\mathcal{F}(p_{vb})$ について, 以下の定理が導出できる.

定理 2. $\mathcal{F}(p_{vb})$ は値を書き下すことができ、以下のようになる.

$$\begin{aligned} \mathcal{F}(p_{vb}) &= KL(p_{vb}(w) \parallel \varphi(w)) - \log C_z, \tag{3.23} \\ KL(p_{vb}(w) \parallel \varphi(w)) &= \sum_{k=1}^K \left\{ (\hat{\phi}_k - \phi) \psi(\hat{\phi}_k) - \log \Gamma(\hat{\phi}_k) + (\hat{\alpha}_k - \alpha) \psi(\hat{\alpha}_k) \right. \\ &\quad - \log \Gamma(\hat{\alpha}_k) + (\hat{\beta}_k - \beta) \psi(\hat{\beta}_k) - \log \Gamma(\hat{\beta}_k) - (\hat{\alpha}_k - \hat{\beta}_k - \alpha - \beta) \psi(\hat{\alpha}_k + \hat{\beta}_k) \\ &\quad \left. + \log \Gamma(\hat{\alpha}_k + \hat{\beta}_k) \right\} + \text{Const.} \tag{3.24} \end{aligned}$$

ここで, Const は学習結果に依らない定数である.

証明. まず, $\mathcal{F}(p_{vb})$ の定義式に式 (3.11) と式 (3.12) を代入すると,

$$\begin{aligned} \mathcal{F}(p_{vb}) &= \sum_{z^n} \int p_{vb}(w) p_{vb}(z^n) \left(\log \frac{p_{vb}(w)}{\varphi(w)} + \log \frac{p_{vb}(z^n)}{p(y^n, z^n | x^n, w)} \right) \\ &= \int p_{vb}(w) \log \frac{p_{vb}(w)}{\varphi(w)} dw \\ &\quad + \sum_{z^n} \int p_{vb}(w) p_{vb}(z^n) \left(E_w [\log p(y^n, z^n | x^n, w)] - \log C_z \right. \\ &\quad \left. - \log p(y^n, z^n | x^n, w) \right) dw \\ &= KL(p_{vb}(w) \parallel \varphi(w)) - \sum_{z^n} \int p_{vb}(w) p_{vb}(z^n) \log C_z dw \\ &= KL(p_{vb}(w) \parallel \varphi(w)) - \log C_z. \end{aligned}$$

ここで, C_z は式 (3.20) によって与えられるため, $p_{vb}(w)$ と $\varphi(w)$ の KL 情報量の計算を行う.

$$\begin{aligned} \log \frac{p_{vb}(w)}{\varphi(w)} &= \sum_{k=1}^K (\hat{\phi}_k - \phi) \log a_k + \sum_{k=1}^K \left\{ (\hat{\alpha}_k - \alpha) \log f_k + (\hat{\beta}_k - \beta) \log(1 - f_k) \right\} \\ &\quad + \left\{ \log \Gamma\left(\sum_{k=1}^K \hat{\phi}_k\right) - \sum_{k=1}^K \log \Gamma(\hat{\phi}_k) \right\} - \left\{ \log \Gamma(K\phi) - K \log \Gamma(\phi) \right\} \\ &\quad + \sum_{k=1}^K \left\{ \log \Gamma(\hat{\alpha}_k + \hat{\beta}_k) - \log \Gamma(\hat{\alpha}_k) - \log \Gamma(\hat{\beta}_k) \right\} \\ &\quad - K \left\{ \log \Gamma(\alpha + \beta) - \log \Gamma(\alpha) - \log \Gamma(\beta) \right\} \end{aligned}$$

であるため、学習結果に依らない定数部分を Const と置き、式 (3.17), (3.18), (3.19) を用いることで最終的に KL 情報量は以下のようになる.

$$\begin{aligned}
KL(p_{vb}(w)||\varphi(w)) &= E_w[\log \frac{p_{vb}(w)}{\varphi(w)}] \\
&= \sum_{k=1}^K \left\{ (\hat{\phi}_k - \phi)\psi(\hat{\phi}_k) - \log \Gamma(\hat{\phi}_k) + (\hat{\alpha}_k - \alpha)\psi(\hat{\alpha}_k) \right. \\
&\quad \left. - \log \Gamma(\hat{\alpha}_k) + (\hat{\beta}_k - \beta)\psi(\hat{\beta}_k) - \log \Gamma(\hat{\beta}_k) - (\hat{\alpha}_k - \hat{\beta}_k - \alpha - \beta)\psi(\hat{\alpha}_k + \hat{\beta}_k) \right. \\
&\quad \left. + \log \Gamma(\hat{\alpha}_k + \hat{\beta}_k) \right\} + \text{Const.}
\end{aligned}$$

以上から、定理を示すことができた. □

この定理の中で、式 (3.23) の導出は、変分ベイズ法の近似分布が式 (3.11), (3.12) の形式で表せるものすべてに対して成り立つものであったため、次章においても利用する. この定理によって $\mathcal{F}(p_{vb})$ の値を計算することができ、この値が小さいほど、学習アルゴリズムにより得られた事後分布 $p_{vb}(w, z^n)$ が式 (3.8) に対してカルバック・ライブラ情報量の意味で良い近似になっているため、本論文では混合数 K のモデル選択や局所解の除去のためこの値を用いる.

3.4 計算機実験

ここまでで、変分ベイズ法による多種粒子 TASEP の学習アルゴリズムの導出ができたので、ここでは、人工データおよび実測データに対して学習アルゴリズムを適用した結果を示す.

3.4.1 人工データへの適用

人工データに対して変分ベイズ法を適用した結果を示す. 人工データを生成するため、本章では真の混合数を 3 として以下の真のパラメータ $w^* = (a_k^*, f_k^*; k = 1, 2, 3)$ を用いた.

$$a^* = (0.33, 0.33, 0.34), f^* = (0.5, 0.7, 0.9).$$

CAにおける各車両の初期位置は等間隔とした．つまり i 番目の車両は L 個あるセルの $\lfloor L \frac{i}{n} \rfloor$ 番目に配置した．車両は $n = 200$ 台用意し，車両の密度が 0.4，つまりセルの個数 $L = 500$ の CA 上で 100 回更新を行った．CA の更新により得られる時空図から試行回数，成功回数の組 (x^n, y^n) の値を計算しそれらを用いて推定を行った．変分ベイズ法におけるハイパーパラメータは $(\phi, \alpha, \beta) = (1, 1, 1)$ を用いた．また，変分ベイズ法の繰り返し回数は 1000 回として，初期値を 500 回変えて $\mathcal{F}(p_{vb})$ が最小になるものを推定結果として用いた．推定が正しく行われたかどうか確認するため，まず，混合数 K に対するモデル選択を汎関数 $\mathcal{F}(p_{vb})$ により行い， $\mathcal{F}(p_{vb})$ の値を最も小さくする混合数 K が真のパラメータの混合数と一致するかということ調べた．次に，その混合数を用いて推定を行い，得られたパラメータの平均値 \hat{w} が真のパラメータに近い値になっているかということ調べた．

以上の実験設定のもとで，混合数 K を 1 から 10 まで変化させた時の汎関数 $\mathcal{F}(p_{vb})$ の値を図 3.1 に示す．横軸が混合数，縦軸が $\mathcal{F}(p_{vb})$ の値を示している．この図から，混合数 $K = 3$ が最小値を与えていることが分かり，したがって $\mathcal{F}(p_{vb})$ によって真の混合数を選択できることを確認した．混合数 $K = 3$ を用いた \hat{w} の値は以下のようになった．

$$\hat{a} = (0.32, 0.34, 0.34), \hat{f} = (0.51, 0.69, 0.90).$$

この結果から，真のパラメータと比較して，推定結果が概ね正確に推定できていることを確認した．

3.4.2 実測データへの適用

人工データに対して正しく推定することが可能であるということを確認したので，次に実測データに適用しどのような結果が得られるか調べた．今回，実測データとして杉山らによって行われた渋滞形成実験で得られた計測データを用いた [71]．この実験は，230 メートルの円周上を車両が走行し，渋滞がどのように形成されるか調べた実験である．実験は 22 台と 23 台の場合でそれぞれ 1 回行われており，それぞれの実験で計測されたデータが公開されている [64]．本論文では，23 台のも

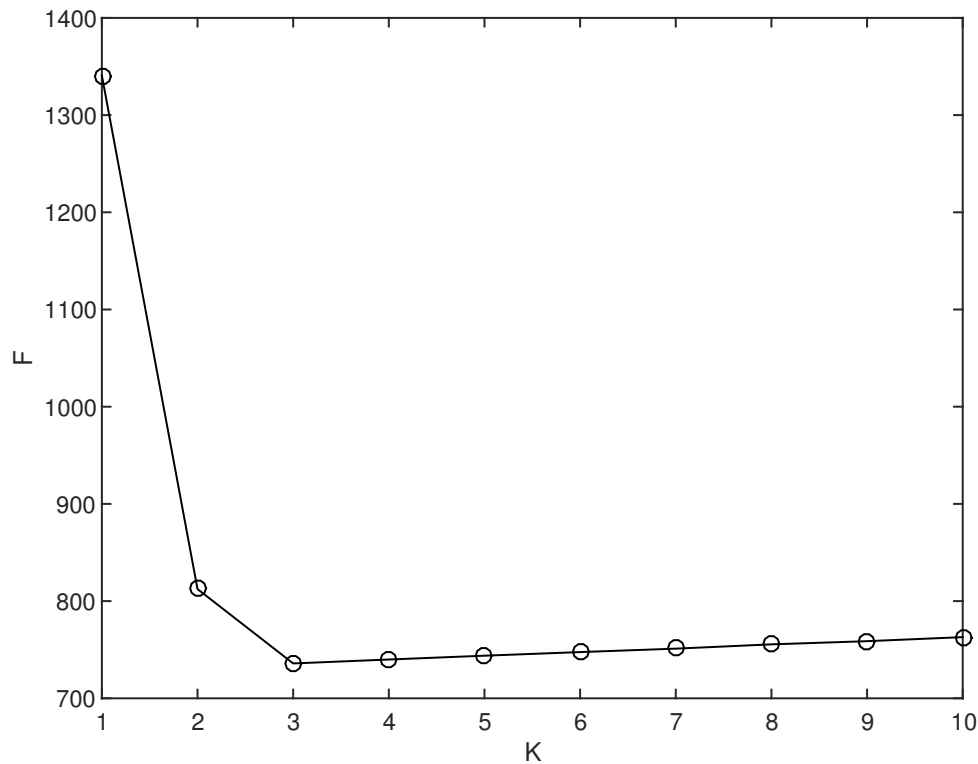


図 3.1: 人工データにおける混合数に対する $F(p_{vb})$ の値

のを用いて推定を行った ($n = 23$). このデータは 0 秒から 500 秒まで 1/3 秒刻みで合計 1501 タイムステップの車両の位置が計測されており, 1000 タイムステップ付近以降で渋滞が発生しているデータである.

本論文では解析を行うため, まず位置データの離散化を行った. 離散化は離散幅が小さすぎる場合, 1 時刻で 2 つ以上前のセルに前進することがあり, 逆に大きすぎる場合, 1 つのセルを同時に 2 つの車両が占有することがある. さらに, 位置データの一部で後退する車両が存在しており, これら 3 つの例外ができるだけ発生しない離散幅を設定することが TASEP で推定を行うために必要である. 今回はこれらの例外が発生しにくい 52 個のセルを用いて離散化を行った. なお, 上述の例外の個数は 52 個のセルを用いた場合 23 個であったのに対し, 例えば 40 個, 70 個のセルを用いると例外はそれぞれ 167 個, 78 個発生した.

この離散化により、周期境界条件を持つ車両の時空図が得られた。また、元のデータは車両番号が分かっているため、どの車両がどのセルを占有しているかを決定することができる。図 3.2(a) に 52 個のセルと 23 台の車両で離散化された円周を示す。円周の外側の数字がセルの番号を表しており、セルの内側の数字はセルの状態を表している。また、セルの値が $s = 0$ である場合車両が存在しないことを表し、 $s = m \in \{1, 2, \dots, 23\}$ の場合は m 番目の車両がいることを表す。さらに、この円周データの時間発展の様子を示すと図 3.2(b) のようになり、これを用いて各車両の「試行回数」と「成功回数」を数えることで推定に用いる車両データ (x^n, y^n) を生成した。さらに本論文では、時間に関して 1501 タイムステップを 6 つに分割し、それぞれについて推定を行った。変分ベイズ法の設定は人工データの場合と同じものを用いて、人工データの場合と同様に混合数 K に関するモデル選択を行い、 $\mathcal{F}(p_{vb})$ を最小にする混合数を用いて推定を行った。

まず、1 から 251 タイムステップの場合の結果を示す。混合数 K を 1 から 23 まで変化させた時の $\mathcal{F}(p_{vb})$ の値を図 3.3 に示す。この図から、 $K = 3$ が最小値を与えることが分かる。 $K = 3$ を用いて得られたパラメータの平均値 \hat{w} の値は表 3.1 のようになった。この結果から、この時刻ではホップ確率が小さい値を持つ車両が全体で一番多く、速い車両が少し含まれているということがわかる。また、実測データを分析する場合、個々の車両が従うホップ確率を定めることは重要である。変分ベイズ法によって得られる隠れ変数の近似分布 $p_{vb}(z^n)$ は、個々の車両に割り当てられるホップ確率の番号の確率を表している。ここでは、 i 番目の車両に k 番目のホップ確率が割り当てられる確率である $\hat{z}_k^{(i)}$ を用いて $\hat{k} = \arg \max_k \hat{z}_k^{(i)}$ により各車両に割り当てられるクラスタ番号を決定する。表 3.2 に各車両に割り当てられたクラスタ番号とそれぞれのクラスタに含まれる車両の数を示す。2 列目が各ホップ確率に属する車両の番号であり、3 列目が属する車両の総数である。例えば、車両番号が 14, 20 である車両は 3 番目のホップ確率に従うことが分かる。他の時刻においても同様の推定を行った。それぞれの時刻における混合数の値に対する $\mathcal{F}(p_{vb})$ 、パラメータの推定結果 \hat{w} 、クラスタ番号の結果を図 3.4 から図 3.8、表 3.3 から表 3.12 に示す。以上の結果から、次のことが確認できた。(1) 多様性を含

んだモデル ($K > 1$) の方が多様性を含まないモデル ($K = 1$) に比べ, $\mathcal{F}(p_{vb})$ の観点で高精度なモデルであり, したがって, モデルの多様性を考慮することで高精度なモデルを表現することができた. (2) 時間に伴って, \hat{w} の値は変化しており, 特に非渋滞時 (1 から 751 タイムステップ), 渋滞手前 (751 から 1001 タイムステップ), 渋滞時 (1001 から 1501 タイムステップ) で大きく値の変化が見られ, 時間に対する非定常性を観測することができた.

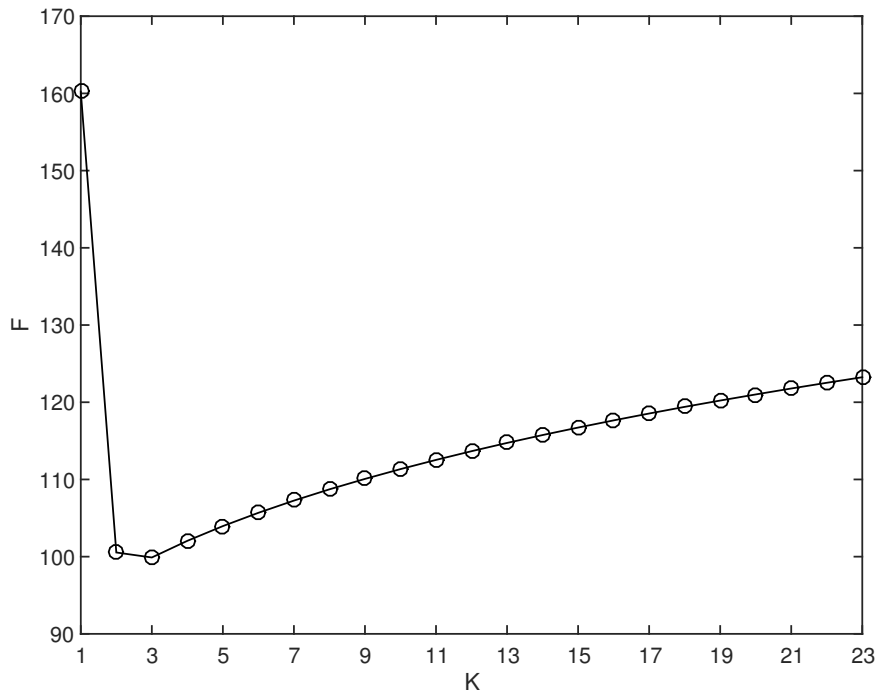


図 3.3: 1 から 251 タイムステップにおける混合数に対する $\mathcal{F}(p_{vb})$ の値 ($K = 3$ が最小)

表 3.2: 1 から 251 タイムステップにおけるクラスタの推定結果

表 3.1: 1 から 251 タイムステップにおけるパラメータの推定結果

	\hat{a}_k	\hat{f}_k
$k = 1$	0.66	0.43
$k = 2$	0.22	0.53
$k = 3$	0.12	0.87

\hat{k}	車両番号	クラスタに含まれる車両数
1	1 4 5 6 7 9 10 11 12 13 16 17 18 21 22 23	16
2	2 3 8 15 19	5
3	14 20	2

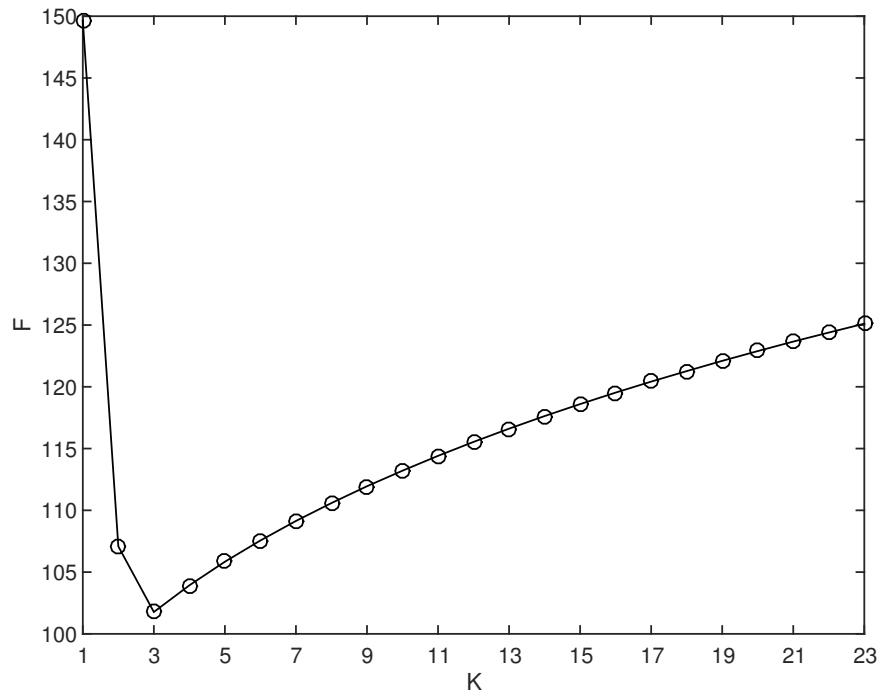


図 3.4: 251 から 501 タイムステップにおける混合数に対する $\mathcal{F}(p_{vb})$ の値 ($K = 3$ が最小)

表 3.4: 251 から 501 タイムステップにおけるクラスタの推定結果

表 3.3: 251 から 501 タイムステップにおけるパラメータの推定結果

	\hat{a}_k	\hat{f}_k
$k = 1$	0.73	0.42
$k = 2$	0.15	0.59
$k = 3$	0.12	0.78

\hat{k}	車両番号	クラスタに含まれる車両数
1	1 3 4 5 6 7 8 9 10 11 12 13 16 17 18 21 22 23	18
2	2 15 19	3
3	14 20	2

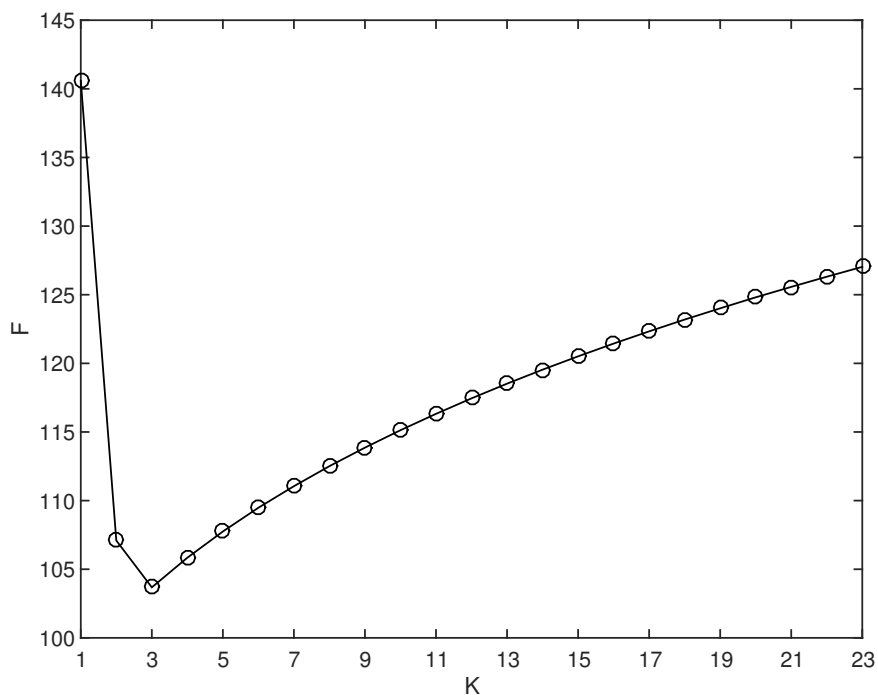


図 3.5: 501 から 751 タイムステップにおける混合数に対する $\mathcal{F}(p_{vb})$ の値 ($K = 3$ が最小)

表 3.6: 501 から 751 タイムステップにおけるクラスタの推定結果

表 3.5: 501 から 751 タイムステップにおけるパラメータの推定結果

	\hat{a}_k	\hat{f}_k
$k = 1$	0.68	0.47
$k = 2$	0.20	0.60
$k = 3$	0.12	0.79

\hat{k}	車両番号	クラスタに含まれる車両数
1	1 4 5 6 7 8 9 10 11 12 13 16 17 18 21 22 23	17
2	2 3 15 19	4
3	14 20	2

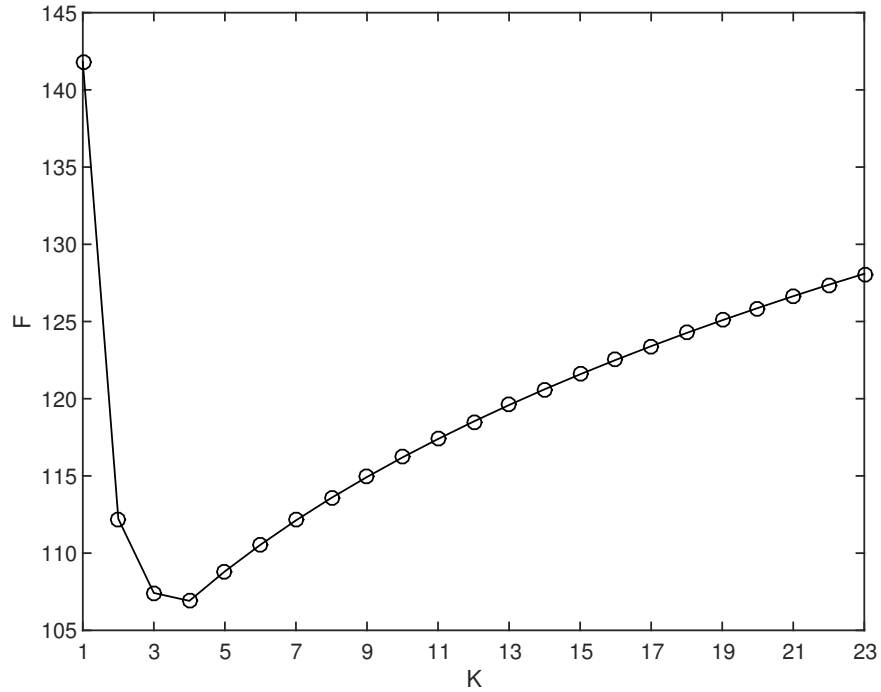


図 3.6: 751 から 1001 タイムステップにおける混合数に対する $\mathcal{F}(p_{vb})$ の値 ($K = 4$ が最小)

表 3.8: 751 から 1001 タイムステップにおけるクラスタの推定結果

表 3.7: 751 から 1001 タイムステップにおけるパラメータの推定結果

	\hat{a}_k	\hat{f}_k
$k = 1$	0.62	0.46
$k = 2$	0.20	0.57
$k = 3$	0.10	0.69
$k = 4$	0.08	0.89

\hat{k}	車両番号	クラスタに含まれる車両数
1	1 4 5 6 7 8	16
	9 10 11 13	
	16 17 18 21	
	22 23	
2	2 3 15 19	4
3	12 14	2
4	20	1

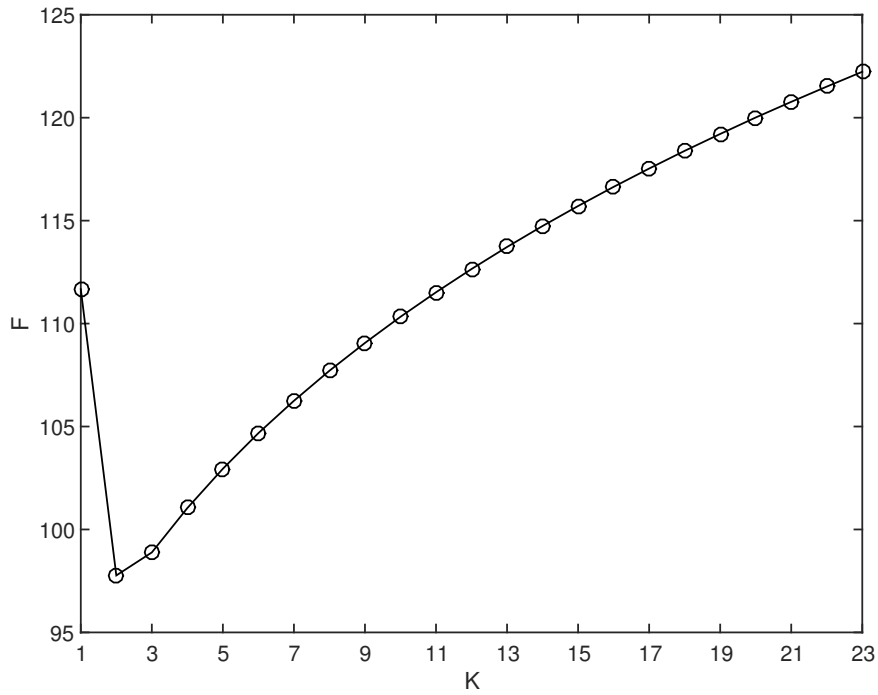


図 3.7: 1001 から 1251 タイムステップにおける混合数に対する $\mathcal{F}(p_{vb})$ の値 ($K = 2$ が最小)

表 3.10: 1001 から 1251 タイムステップにおけるクラスタの推定結果

表 3.9: 1001 から 1251 タイムステップにおけるパラメータの推定結果

	\hat{a}_k	\hat{f}_k
$k = 1$	0.80	0.39
$k = 2$	0.20	0.55

\hat{k}	車両番号	クラスタに含まれる車両数
1	1 2 3 4 5 6	19
	7 8 9 10 11	
	12 13 14 17	
	18 21 22 23	
2	15 16 19 20	4

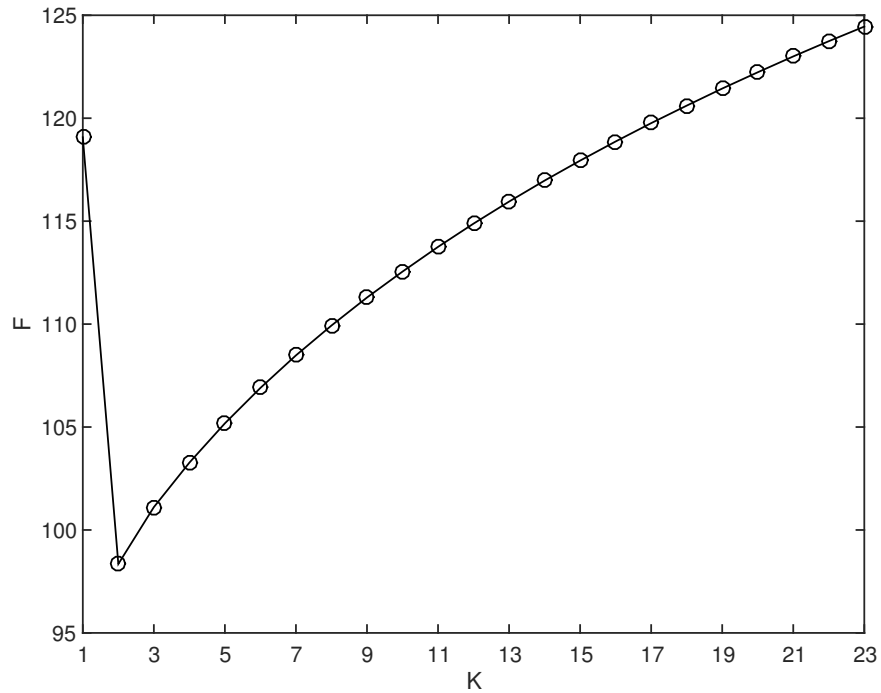


図 3.8: 1251 から 1501 タイムステップにおける混合数に対する $\mathcal{F}(p_{vb})$ の値 ($K = 2$ が最小)

表 3.12: 1251 から 1501 タイムステップにおけるクラスタの推定結果

表 3.11: 1251 から 1501 タイムステップにおけるパラメータの推定結果

	\hat{a}_k	\hat{f}_k
$k = 1$	0.84	0.42
$k = 2$	0.16	0.63

\hat{k}	車両番号	クラスタに含まれる車両数
1	1 2 3 4 5 6 7 8 9 10 11 12 13 16 17 18 19 21 22 23	20
2	14 15 20	3

3.5 考察

ここでは、多種粒子 TASEP の推定で用いた学習データが推定で得られたパラメータと整合性の取れた結果であるかということを議論し、CA の端と端がつながっていない開放系の CA の場合の多種粒子 TASEP の推定方法について議論する。

3.5.1 学習データと推定結果の整合性の確認

学習データと推定結果の整合性について議論するため、ホップ確率が低いクラスタに属する車両と高いクラスタに属する車両の時空図上での動きについて議論する。図 3.9 に 100 から 200 タイムステップにおける時空図を示す。縦軸がタイムステップ、横軸がセルの位置を表しており、図中の黒い部分に車両が存在していることを表している。(a) は、表 3.2 の中で、ホップ確率が低いクラスタに属する車両番号として車両番号 1 と 7 を選び、2 つの車両の軌跡を濃い色で強調した時空図である。また、ホップ確率が高いクラスタに属する車両の番号は 14 と 20 であり、これらの車両の軌跡を濃い色で強調した時空図が (b) である。この 2 つの図から、ホップ確率が高いクラスタに属する車両は低いクラスタの車両に比べ、前の車両の軌跡に接近した進み方をしており、前進しやすい傾向があることを確認できた。また、 i 番目の車両のホップ確率 $\hat{f}^{(i)}$ は $\hat{f}^{(i)} = \frac{y^{(i)}}{x^{(i)}}$ によって与えられており、車両番号 1, 7, 14, 20 の車両の 100 から 200 タイムステップのホップ確率は次のようになった。

$$\hat{f}^{(1)} = 0.41, \hat{f}^{(7)} = 0.39, \hat{f}^{(14)} = 0.89, \hat{f}^{(20)} = 0.84.$$

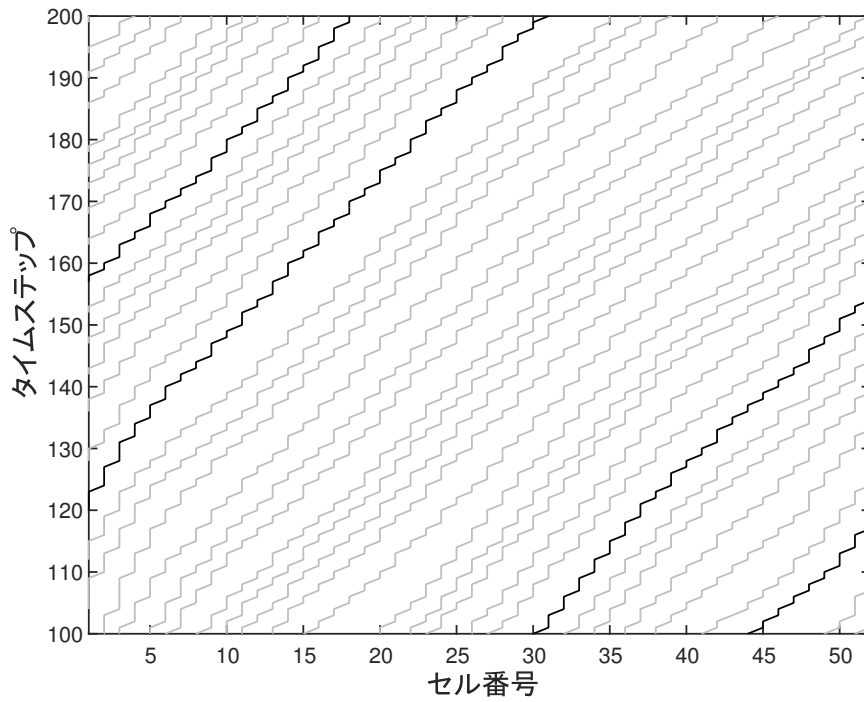
これより、各車両のホップ確率が近い者同士が同じクラスタに属していることも確認できた。したがって、今回の推定結果は、時空図上の車両の動きと各車両のホップ確率の観点で整合性の取れたものであることが分かった。

3.5.2 開放系の CA における多種粒子 TASEP の推定

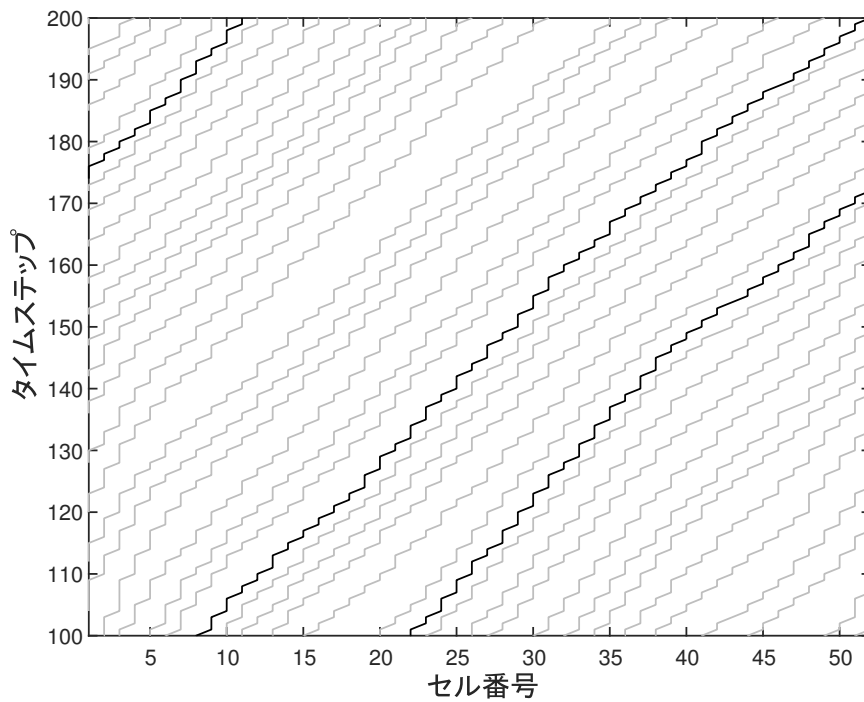
次に，CA の端と端がつながっていない開放系の CA モデルにおける多種粒子 TASEP のパラメータ推定方法について考える．セルの全長が L の場合，各車両の成功回数は固定であり，成功回数 $y = L$ である．セルの端に到達するまでに要した試行回数 x は各車両に割り当てられるホップ確率に依存して異なるため，3.2 節と同じパラメータ表記を用いると試行回数の分布 $p(x|L, w)$ は

$$p(x|L, w) = \sum_{k=1}^K a_k \binom{x}{L} f_k^L (1 - f_k)^{x-L}$$

であり，負の 2 項分布が比率 a_k で混合した分布で表現することができる．この式で L は与えられおり， x は時空図から計算することができ，またパラメータに注目すると，式 (3.5) と同じ形である．したがって，開放系の CA モデルは車両の流入，流出の確率を考慮する必要があるが，それを除くと今回考えた周期境界条件を持つ CA の推定と同様の推定アルゴリズムにより推定可能である．ゆえに，実問題でしばしば想定されるセルの端と端がつながっていない場合においても本論文の推定方法は有用である [93]．



(a) ホップ確率が低い車両 (車両番号 1, 7) に注目した時空図



(b) ホップ確率が高い車両 (車両番号 14, 20) に注目した時空図

図 3.9: 100 から 200 タイムステップにおける時空図

第4章 ZRPの変分ベイズ学習

本章では，TASEPの場合と同じようにZRPの確率モデルである統計的表現について4.1節で全ての車両が同じOV関数に従う(通常の)ZRPの場合について紹介し，4.2節で複数のOV関数が混在する多種粒子ZRPの場合について述べる．その後，4.3節で多種粒子ZRPに変分ベイズ法を適用した時に得られる学習アルゴリズムの導出を行い，4.4節で人工データ，実測データに適用した結果を示す．最後に4.5節で推定結果の確からしさと多種粒子TASEP，多種粒子ZRPの比較に関する考察を行う [94].

4.1 ZRPの統計的表現

ZRPではOV関数 f_d がパラメータである．ある時刻で車間距離 $d = 0$ の場合，体積排除効果により確率1で同じセルに留まるため $f_0 = 0$ である．車間距離 d が0より大きい場合，確率 f_d で「前に進む」という事象が発生し，確率 $1 - f_d$ で「同じセルに残る」という事象が発生する．これらの事象は前の時刻の事象に依らず独立に生起するため，ベルヌーイ試行である．「前に進む」という事象を「成功」，「同じセルに残る」という事象を「失敗」とみなすと，車間距離 d における「成功」の回数 y_d の分布は車間距離 d における試行回数（「成功」と「失敗」の回数の合計） x_d が与えられたもとのOV関数 f_d を用いて2項分布で表すことができる．

$$p(y_d|x_d) = \binom{x_d}{y_d} f_d^{y_d} (1 - f_d)^{x_d - y_d}. \quad (4.1)$$

さらに各車間距離における成功回数がお互いに独立に発生するため，車間距離の最大値を M とすると，同時分布 $p(y_1, \dots, y_M | x_1, \dots, x_M)$ は各車間距離ごとの2項分

布の積で表すことができ,

$$p(y_1, \dots, y_M | x_1, \dots, x_M) = \prod_{j=1}^M \binom{x_j}{y_j} f_j^{y_j} (1 - f_j)^{x_j - y_j} \quad (4.2)$$

によって表すことができる. これが全ての車両が同じ OV 関数に従う ZRP の統計的表現である. 本章では, 以後 M 個の組 (y_1, \dots, y_M) と (x_1, \dots, x_M) を y, x でそれぞれ表記する. なお, $y_{sum} = \sum_{j=1}^M y_j, x_{sum} = \sum_{j=1}^M x_j$ とすると, $f_1 = f_2 = \dots = f_M = f$ である時, y_{sum} の確率分布 $p(y_{sum} | x_{sum})$ は TASEP の統計的表現と一致するため, ZRP の統計的表現は TASEP の統計的表現を拡張したものであることがわかる.

4.2 多種粒子 ZRP の統計的表現

多種粒子 ZRP では, K 種類の OV 関数 f_{1d}, \dots, f_{Kd} が比率 a_1, \dots, a_K で割り当てられる. 多種粒子 TASEP の場合と同じように, 割り当てられる OV 関数の番号を表現するため, 隠れ変数 $Z \in \mathcal{C}$ を導入すると $p(z)$ が式 (3.2) で表すことができ, z, x が与えられたもとの成功回数 y の分布 $p(y|x, z)$ は

$$p(y|x, z) = \prod_{k=1}^K \left\{ \prod_{j=1}^M \binom{x_j}{y_j} f_{kj}^{y_j} (1 - f_{kj})^{x_j - y_j} \right\}^{z_k}$$

となる. また, (y, z) の同時分布 $p(y, z|x)$ は

$$p(y, z|x) = p(y|x, z)p(z) = \prod_{k=1}^K \left\{ a_k \prod_{j=1}^M \binom{x_j}{y_j} f_{kj}^{y_j} (1 - f_{kj})^{x_j - y_j} \right\}^{z_k} \quad (4.3)$$

となり, 隠れ変数 z を周辺化することで多種粒子 ZRP の統計的表現は

$$p(y|x) = \sum_{k=1}^K a_k \prod_{j=1}^M \binom{x_j}{y_j} f_{kj}^{y_j} (1 - f_{kj})^{x_j - y_j} \quad (4.4)$$

で表すことができる. 多種粒子 ZRP においてパラメータは, 比率 a_1, \dots, a_K と OV 関数 $(f_{kj}; k = 1, \dots, K, j = 1, \dots, M)$ である. これらをまとめて, $w = (a_k, f_{kj}; k = 1, \dots, K, j = 1, \dots, M)$ で表す. 本章では, パラメータを明記する場合, 以後式 (4.4) は $p(y|x, w)$ で表す.

4.3 多種粒子 ZRP の変分ベイズ法の学習アルゴリズム

ここでは、変分ベイズ法による多種粒子 ZRP のパラメータ推定について考える。 n 個の x, y, z の集合を $x^n = \{x^{(1)}, \dots, x^{(n)}\}$, $y^n = \{y^{(1)}, \dots, y^{(n)}\}$, $z^n = \{z^{(1)}, \dots, z^{(n)}\}$ で表すと、3章と同じように、隠れ変数とデータの同時分布を $p(y^n, z^n | x^n, w)$ が定義することができ、事前分布 $\varphi(w)$ を用いて式 (3.11), (3.12) から変分ベイズ法の更新式を求めることができる。

まず、事前分布 $\varphi(w)$ として以下のものを用いる。

$$\varphi(w) = \text{Dir}(a | \{\phi_k\}_{k=1}^K) \prod_{k=1}^K \prod_{j=1}^M \text{Beta}(f_{kj} | \alpha_{kj}, \beta_{kj}). \quad (4.5)$$

ここで $(\phi_k, \alpha_{kj}, \beta_{kj}; k = 1, \dots, K, j = 1, \dots, M)$ はハイパーパラメータであり、3章の事前分布と同じように全ての k, j について $\phi_k = \phi > 0, \alpha_{kj} = \alpha > 0, \beta_{kj} = \beta > 0$ とする。また、 $p(y^n, z^n | x^n, w)$ は3章の式 (3.7) の形で与えられるため、式 (4.3) を用いて同時分布の対数尤度は

$$\begin{aligned} \log p(y^n, z^n | x^n, w) = & \sum_{k=1}^K \sum_{i=1}^n z_k^{(i)} \left[\log a_k + \sum_{j=1}^M \left\{ y_j^{(i)} \log f_{kj} \right. \right. \\ & \left. \left. + (x_j^{(i)} - y_j^{(i)}) \log(1 - f_{kj}) + \log \binom{x_j^{(i)}}{y_j^{(i)}} \right\} \right] \end{aligned}$$

となる。これと $\hat{z}_k^{(i)} = E_z[z_k^{(i)}]$ を用いると、多種粒子 ZRP における変分ベイズ法によって得られるパラメータの分布 $p_{vb}(w)$ は以下ようになる。

$$p_{vb}(w) = \text{Dir}(a | \{\hat{\phi}_k\}_{k=1}^K) \prod_{k=1}^K \prod_{j=1}^M \text{Beta}(f_{kj} | \hat{\alpha}_{kj}, \hat{\beta}_{kj}), \quad (4.6)$$

$$\hat{\phi}_k = \hat{z}_k^{(i)} + \phi, \quad (4.7)$$

$$\hat{\alpha}_{kj} = \sum_{i=1}^n y_j^{(i)} \hat{z}_k^{(i)} + \alpha, \quad (4.8)$$

$$\hat{\beta}_{kj} = \sum_{i=1}^n (x_j^{(i)} - y_j^{(i)}) \hat{z}_k^{(i)} + \beta. \quad (4.9)$$

また、パラメータの対数期待値が

$$E_w[\log a_k] = \psi(\hat{\phi}_k) - \psi\left(\sum_{k=1}^K \hat{\phi}_k\right), \quad (4.10)$$

$$E_w[\log f_{kj}] = \psi(\hat{\alpha}_{kj}) - \psi(\hat{\alpha}_{kj} + \hat{\beta}_{kj}), \quad (4.11)$$

$$E_w[\log(1 - f_k)] = \psi(\hat{\beta}_{kj}) - \psi(\hat{\alpha}_{kj} + \hat{\beta}_{kj}) \quad (4.12)$$

であることに注意し、得られた $(\hat{\phi}_k, \hat{\alpha}_{kj}, \hat{\beta}_{kj})$ を用いて、多種粒子 ZRP における隠れ変数の分布 $p_{vb}(z^n)$ は

$$\begin{aligned} p_{vb}(z^n) &= \frac{1}{C_z} \prod_{i=1}^n \prod_{k=1}^K \exp(z_k^{(i)} L_{ik}), \\ L_{ik} &= \psi(\hat{\phi}_k) + \sum_{j=1}^M \left\{ y_j^{(i)} \psi(\hat{\alpha}_{kj}) + (x_j^{(i)} - y_j^{(i)}) \psi(\hat{\beta}_{kj}) - x_j^{(i)} \psi(\hat{\alpha}_{kj} + \hat{\beta}_{kj}) \right\}, \\ C_z &= \prod_{i=1}^n \sum_{k=1}^K \exp(L_{ik}), \end{aligned} \quad (4.13)$$

となる。 $p_{vb}(z^n)$ を用いて

$$\hat{z}_k^{(i)} = \frac{\exp(L_{ik})}{\sum_{l=1}^K \exp(L_{il})} \quad (4.14)$$

の計算を行い、これを繰り返すことでパラメータの分布と隠れ変数の分布を求める。以上をまとめると、学習アルゴリズムは次のようになる。

アルゴリズム 2 (多種粒子 ZRP の変分ベイズ法のアルゴリズム).

1. $\{\hat{z}_k^{(i)}; i = 1, \dots, n, k = 1, \dots, K\}$ を初期化する。

2. M-ステップ: $p_{vb}(w)$ の更新

$$\begin{aligned} \hat{\phi}_k &\leftarrow \sum_{i=1}^n \hat{z}_k^{(i)} + \phi, \\ \hat{\alpha}_{kj} &\leftarrow \sum_{i=1}^n y_j^{(i)} \hat{z}_k^{(i)} + \alpha, \\ \hat{\beta}_{kj} &\leftarrow \sum_{i=1}^n (x_j^{(i)} - y_j^{(i)}) \hat{z}_k^{(i)} + \beta. \end{aligned}$$

3. E-ステップ: $p_{vb}(z^n)$ の更新

$$L_{ik} \leftarrow \psi(\hat{\phi}_k) + \sum_{j=1}^M \left\{ y_j^{(i)} \psi(\hat{\alpha}_{kj}) + (x_j^{(i)} - y_j^{(i)}) \psi(\hat{\beta}_{kj}) - x_j^{(i)} \psi(\hat{\alpha}_{kj} + \hat{\beta}_{kj}) \right\},$$

$$\hat{z}_k^{(i)} \leftarrow \frac{\exp(L_{ik})}{\sum_{l=1}^K \exp(L_{il})}.$$

4. E-ステップと M-ステップを交互に繰り返し更新する.

多種粒子 ZRP におけるパラメータの平均値 $\hat{w} = (\hat{a}_k, \hat{f}_{kj}; k = 1, \dots, K, j = 1, \dots, M)$ は

$$\hat{a}_k = \frac{\hat{\phi}_k}{\sum_{l=1}^K \hat{\phi}_l}, \hat{f}_{kj} = \frac{\hat{\alpha}_{kj}}{\hat{\alpha}_{kj} + \hat{\beta}_{kj}} \quad (4.15)$$

となる. また, 予測分布 $p_{vb}(y|x, x^n, y^n)$ は

$$\begin{aligned} p_{vb}(y|x, x^n, y^n) &= \sum_{k=1}^K E_w[a_k] \prod_{j=1}^M \binom{x_j}{y_j} E_w[f_{kj}^{y_j} (1 - f_{kj})^{x_j - y_j}] \\ &= \sum_{k=1}^K \hat{a}_k \prod_{j=1}^M \binom{x_j}{y_j} \frac{\text{Beta}(\hat{\alpha}_{kj} + y_j, \hat{\beta}_{kj} + (x_j - y_j))}{\text{Beta}(\hat{\alpha}_{kj}, \hat{\beta}_{kj})} \end{aligned} \quad (4.16)$$

である.

多種粒子 TASEP の場合と同じく, $\mathcal{F}(p_{vb})$ について以下の定理を導出することができる.

定理 3. $\mathcal{F}(p_{vb})$ は値を書き下すことができ, 以下のようになる.

$$\mathcal{F}(p_{vb}) = KL(p_{vb}(w) || \varphi(w)) - \log C_z, \quad (4.17)$$

$$\begin{aligned} KL(p_{vb}(w) || \varphi(w)) &= \sum_{k=1}^K \left[(\hat{\phi}_k - \phi) \psi(\hat{\phi}_k) - \log \Gamma(\hat{\phi}_k) + \sum_{j=1}^M \left\{ (\hat{\alpha}_{kj} - \alpha) \psi(\hat{\alpha}_{kj}) \right. \right. \\ &\quad - \log \Gamma(\hat{\alpha}_{kj}) + (\hat{\beta}_{kj} - \beta) \psi(\hat{\beta}_{kj}) - \log \Gamma(\hat{\beta}_{kj}) - (\hat{\alpha}_{kj} - \hat{\beta}_{kj} - \alpha - \beta) \psi(\hat{\alpha}_{kj} + \hat{\beta}_{kj}) \\ &\quad \left. \left. + \log \Gamma(\hat{\alpha}_{kj} + \hat{\beta}_{kj}) \right\} \right] + \text{Const} \end{aligned} \quad (4.18)$$

ここで, C_z は式 (4.13) であり, Const は学習結果に依らない定数である.

証明. 多種粒子 ZRP が多種粒子 TASEP の場合と同じく変分ベイズ法の近似分布が式 (3.11), (3.12) の形であるため, $\mathcal{F}(p_{vb})$ は式 (4.17) のように分解できる. また, $\log \frac{p_{vb}(w)}{\varphi(w)}$ は式 (4.6) と式 (4.5) を用いて

$$\begin{aligned} \log \frac{p_{vb}(w)}{\varphi(w)} &= \sum_{k=1}^K (\hat{\phi}_k - \phi) \log a_k + \sum_{k=1}^K \sum_{j=1}^M \left\{ (\hat{\alpha}_{kj} - \alpha) \log f_k + (\hat{\beta}_{kj} - \beta) \log(1 - f_{kj}) \right\} \\ &\quad + \left\{ \log \Gamma \left(\sum_{k=1}^K \hat{\phi}_k \right) - \sum_{k=1}^K \log \Gamma(\hat{\phi}_k) \right\} - \left\{ \log \Gamma(K\phi) - K \log \Gamma(\phi) \right\} \\ &\quad + \sum_{k=1}^K \sum_{j=1}^M \left\{ \log \Gamma(\hat{\alpha}_{kj} + \hat{\beta}_{kj}) - \log \Gamma(\hat{\alpha}_{kj}) - \log \Gamma(\hat{\beta}_{kj}) \right\} \\ &\quad - KM \left\{ \log \Gamma(\alpha + \beta) - \log \Gamma(\alpha) - \log \Gamma(\beta) \right\} \end{aligned}$$

である. 学習結果に依らない定数部分を Const と置き, 式 (4.10), (4.11), (4.12) を用いて $p_{vb}(w)$ で平均を取ると, 式 (4.18) が得られる. この結果と C_z が多種粒子 ZRP の場合, 式 (4.13) で表せることから, 定理を示すことができた. \square

4.4 計算機実験

ここまでで変分ベイズ法による多種粒子 ZRP の学習アルゴリズムの導出ができたので多種粒子 TASEP の場合と同じく, 人工データ, 実測データに適用した結果を示す.

4.4.1 人工データへの適用

人工データを生成するため, 本章では真の混合数が 4 であり, 真のパラメータ $w^* = (a_k^*, f_{kj}^*; k = 1, 2, 3, 4, j = 1, 2, 3, 4)$ が図 4.1, 表 4.1 であるものを用いた. 図 4.1 が各 OV 関数 f_{kj}^* の値であり, 表 4.1 が各 OV 関数の混合比 a_k^* である. つまり, 割り当てられる OV 関数は等確率で, 目標速度 (OV 関数の最大値) も車間も詰めにくい車両 ($k = 1$), 目標速度は遅いが車間は詰めやすい車両 ($k = 2$), 目標速度は速いが車間を詰めにくい車両 ($k = 3$), 目標速度が速く車間を詰めやすい車両

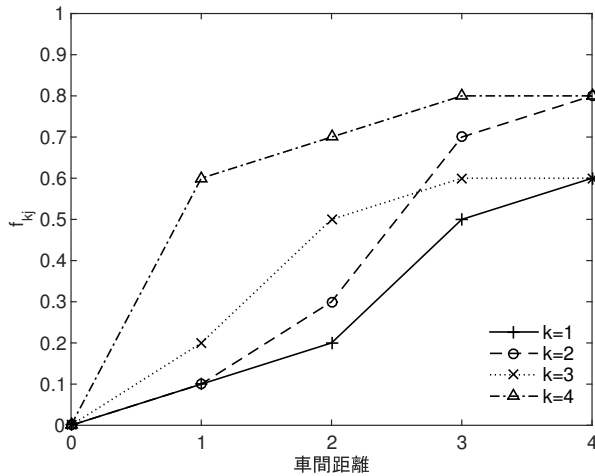


表 4.1: 人工データを生成する真の混合比 a_k^*

	a_k^*
$k = 1$	0.25
$k = 2$	0.25
$k = 3$	0.25
$k = 4$	0.25

図 4.1: 人工データを生成する真の OV 関数 f_{kj}^*

($k = 4$) が CA 上を走行する状況を想定している。また、CA における各車両の初期位置は等間隔とした。つまり i 番目の車両は L 個あるセルの $\lfloor L \frac{i}{n} \rfloor$ 番目に配置した。また、車間距離が 4 より大きい場合、車間距離が 4 のものとして扱った。車両は $n = 200$ 台用意し、車両の密度が 0.4、つまりセルの個数 $L = 500$ の CA 上で 100 回更新を行った。CA の更新により得られる時空図から各車間距離における試行回数、成功回数の組 (x^n, y^n) の値を計算しそれらを用いて推定を行った。変分ベイズ法のハイパーパラメータは $(\phi, \alpha, \beta) = (1, 1, 1)$ として、多種粒子 TASEP の場合と同じく、1000 回繰り返し更新を行い、初期値を 500 回変え $\mathcal{F}(p_{vb})$ を最小にするものを推定結果として用いた。

まず、混合数 K を 1 から 10 まで変化させたときの $\mathcal{F}(p_{vb})$ の値を図 4.2 に示す。横軸が混合数、縦軸が $\mathcal{F}(p_{vb})$ の値である。この図から、 $K = 4$ が最小値を与えており、 $\mathcal{F}(p_{vb})$ により真の混合数を選択できることを確認した。次に混合数 $K = 4$ を用いた \hat{w} の値について図 4.3 に各 OV 関数の平均値 \hat{f}_{kj} を、表 4.2 に混合比の平均値 \hat{a}_k を示した。真のパラメータと比較すると、 $k = 4$ の OV 関数の車間距離 4 の値が真の OV 関数と大きく異なっていることがわかる。これは、この部分を推定するためのデータが足りないことが原因であると考えられ、実測データにおいても同様の問題が発生しており 4.5 節で議論する。以上の結果から、 $\mathcal{F}(p_{vb})$ のモ

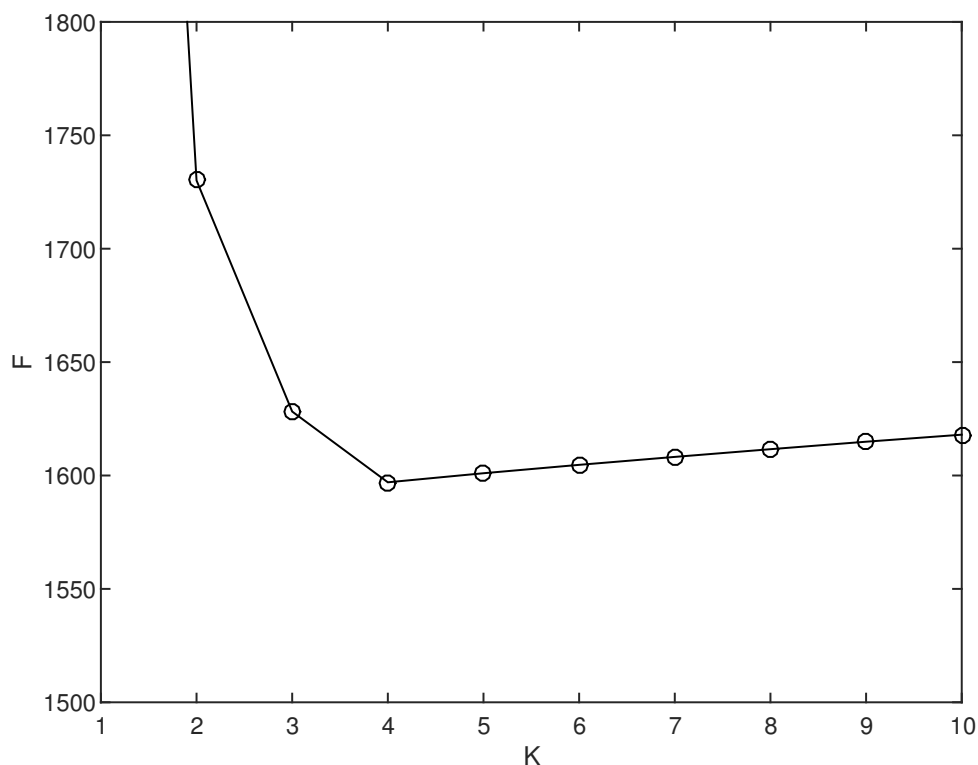


図 4.2: 人工データにおける混合数に対する $\mathcal{F}(p_{vb})$ の値

モデル選択により正しくモデルを選ぶことができ、正しいモデルで推定を行うことでデータが足りない部分を除いて概ね正確にパラメータを推定できることを確認した。

4.4.2 実測データへの適用

3章と同じように杉山らによって行われた実測データを用いた [64]。本章においても、 $n = 23$ 台の計測データを用いて時間を 6 つに、円周を 52 個のセルに分割し、車両の位置を離散化することで ZRP で用いる車両データ (x^n, y^n) を生成した。車間距離は 4 より大きい場合、車間距離が 4 のものとして扱った。変分ベイズ法の設定は人工データの場合と同じものを用いて、混合数 K に関するモデル選択を $K = 1$ から 23 まで行い、 $\mathcal{F}(p_{vb})$ を最小にする混合数を用いて推定を行った。

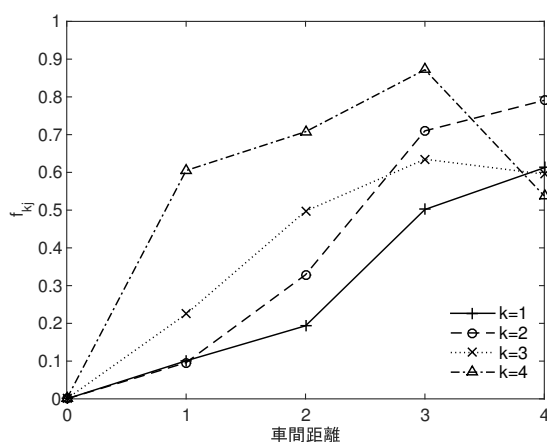


図 4.3: 人工データに対する OV 関数の推定結果

$$\hat{f}_{kj}$$

表 4.2: 人工データに対する混合比の推定結果 \hat{a}_k

	\hat{a}_k
$k = 1$	0.23
$k = 2$	0.27
$k = 3$	0.24
$k = 4$	0.26

図 4.4 から図 4.15, 表 4.3 から表 4.14 は 1 から 251 タイムステップから 1251 から 1501 タイムステップまでのそれぞれの時刻における推定結果を表す. それぞれ図 4.4, 4.6, 4.8, 4.10, 4.12, 4.14 が各時刻における混合数 K に対する $\mathcal{F}(p_{vb})$ の値であり, 図 4.5, 4.7, 4.9, 4.11, 4.13, 4.15 がモデル選択で選ばれた混合数 K を用いた OV 関数の推定結果 \hat{f}_{kj} であり, 凡例にそれぞれの OV 関数の番号を示している. 表 4.3, 4.5, 4.7, 4.9, 4.11, 4.13 が各 OV 関数の混合比の推定結果 \hat{a}_k であり, 表 4.4, 4.6, 4.8, 4.10, 4.12, 4.14 が各車両に割り当てられる OV 関数の結果を示している.

以上から, 各時刻で単一の OV 関数で推定を行う場合に比べ, 複数の OV 関数を用いた方が $\mathcal{F}(p_{vb})$ に関して良い学習結果を与えることが分かり, 各時刻における非定常性も確認することができた. また, 人工データの場合と同じように, 例えば図 4.5 の $k = 6$ における車間距離 2, 3, 4 のように OV 関数の値が大きく減少している部分を確認することができた.

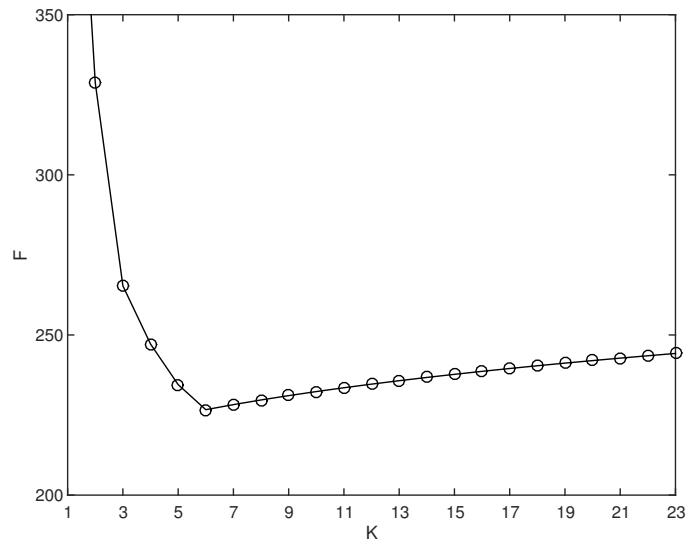


図 4.4: 1 から 251 タイムステップにおける混合数に対する $\mathcal{F}(p_{ob})$ の値 ($K = 6$ が最小)

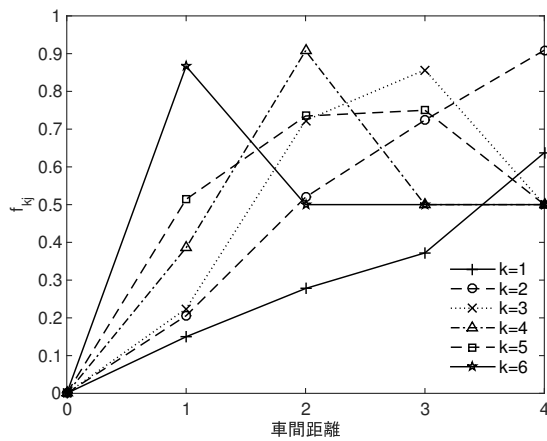


図 4.5: 1 から 251 タイムステップにおける OV 関数の推定結果 \hat{f}_{kj}

表 4.3: 1 から 251 タイムステップにおける混合比の推定結果 \hat{a}_k

	\hat{a}_k
$k = 1$	0.07
$k = 2$	0.24
$k = 3$	0.21
$k = 4$	0.21
$k = 5$	0.17
$k = 6$	0.10

表 4.4: 1 から 251 タイムステップにおけるクラスタの推定結果

\hat{k}	車両番号	クラスタに含まれる車両数
1	7	1
2	4 5 10 11 13 18	6
3	1 6 17 21 23	5
4	8 9 12 16 22	5
5	2 3 15 19	4
6	14 20	2

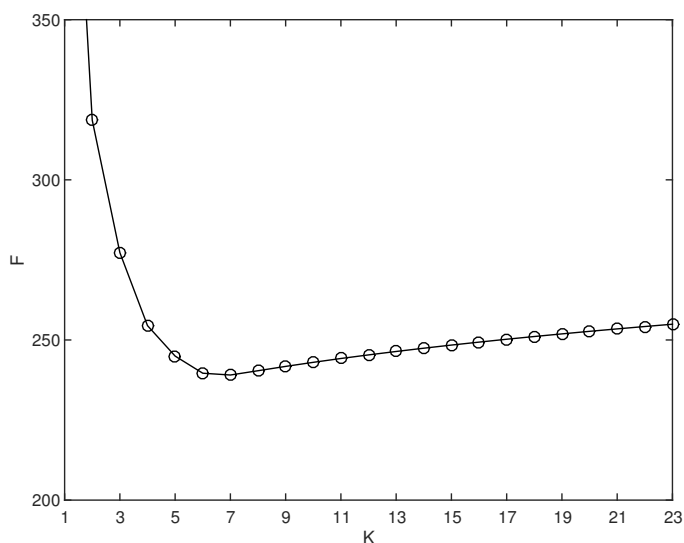


図 4.6: 251 から 501 タイムステップにおける混合数に対する $\mathcal{F}(p_{vb})$ の値 ($K = 7$ が最小)

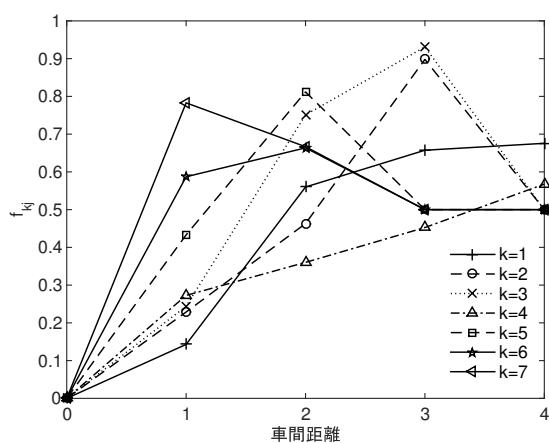


図 4.7: 251 から 501 タイムステップにおける OV 関数の推定結果 \hat{f}_{kj}

表 4.5: 251 から 501 タイムステップにおける混合比の推定結果 \hat{a}_k

	\hat{a}_k
$k = 1$	0.13
$k = 2$	0.10
$k = 3$	0.27
$k = 4$	0.10
$k = 5$	0.17
$k = 6$	0.13
$k = 7$	0.10

表 4.6: 251 から 501 タイムステップにおけるクラスターの推定結果

\hat{k}	車両番号	クラスターに含まれる車両数
1	5 11 18	3
2	4 13	2
3	1 6 8 9 17 21 23	7
4	7 10	2
5	3 12 16 22	4
6	2 15 19	3
7	14 20	2

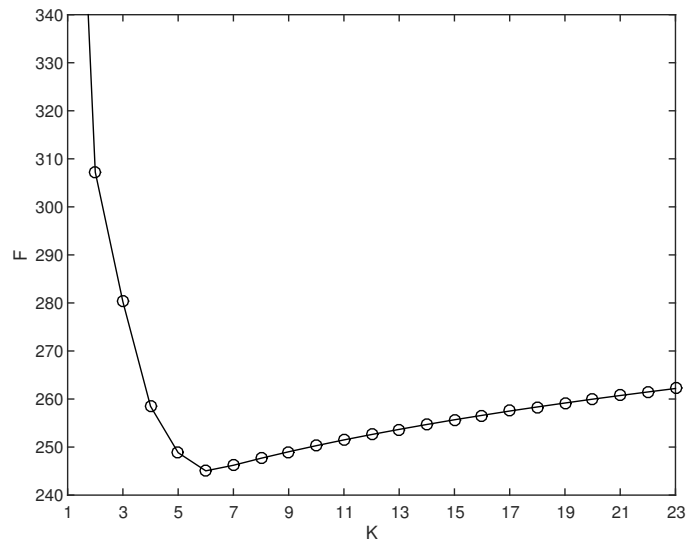


図 4.8: 501 から 751 タイムステップにおける混合数に対する $\mathcal{F}(p_{vb})$ の値 ($K = 6$ が最小)

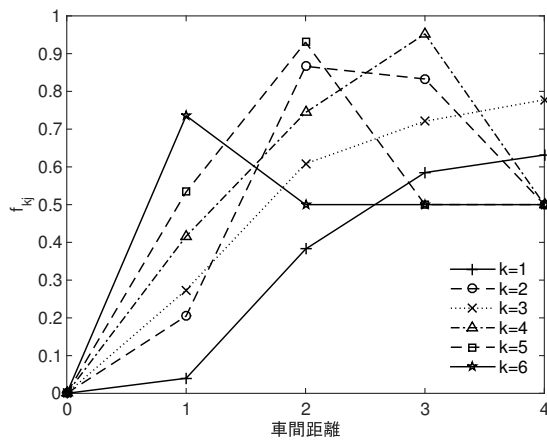


図 4.9: 501 から 751 タイムステップにおける OV 関数の推定結果 \hat{f}_{kj}

表 4.7: 501 から 751 タイムステップにおける混合比の推定結果 \hat{a}_k

	\hat{a}_k
$k = 1$	0.07
$k = 2$	0.10
$k = 3$	0.35
$k = 4$	0.17
$k = 5$	0.17
$k = 6$	0.14

表 4.8: 501 から 751 タイムステップにおけるクラスターの推定結果

\hat{k}	車両番号	クラスターに含まれる車両数
1	13	1
2	6 9	2
3	1 4 5 7 10 11 17 18 21	9
4	8 16 22 23	4
5	2 3 12 15	4
6	14 19 20	3

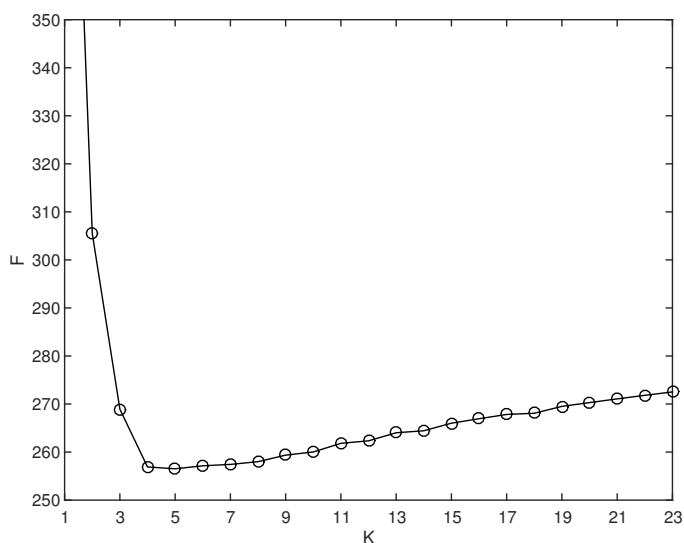


図 4.10: 751 から 1001 タイムステップにおける混合数に対する $F(p_{vb})$ の値 ($K = 5$ が最小)

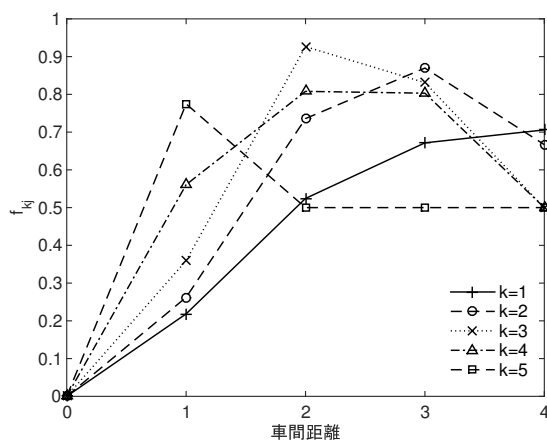


表 4.9: 751 から 1001 タイムステップにおける混合比の推定結果 \hat{a}_k

	\hat{a}_k
$k = 1$	0.29
$k = 2$	0.25
$k = 3$	0.14
$k = 4$	0.22
$k = 5$	0.10

図 4.11: 751 から 1001 タイムステップにおける OV 関数の推定結果 \hat{f}_{kj}

表 4.10: 751 から 1001 タイムステップにおけるクラスターの推定結果

\hat{k}	車両番号	クラスターに含まれる車両数
1	1 7 8 10 11 13 21	7
2	4 5 6 17 22 23	6
3	9 16 18	3
4	2 3 12 15 19	5
5	14 20	2

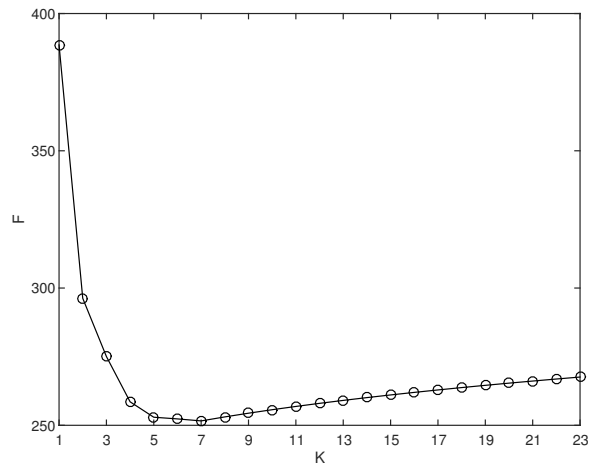


図 4.12: 1001 から 1251 タイムステップにおける混合数に対する $\mathcal{F}(p_{vb})$ の値 ($K = 7$ が最小)

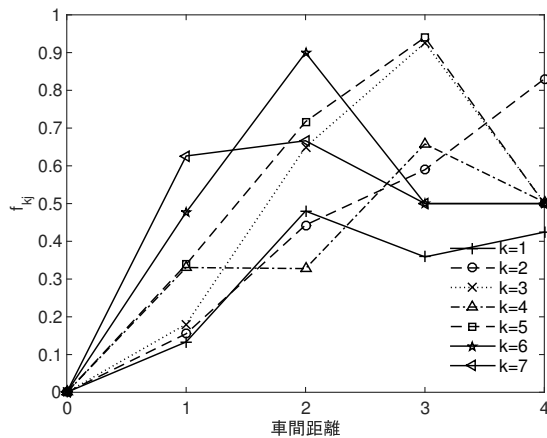


図 4.13: 1001 から 1251 タイムステップにおける OV 関数の推定結果 \hat{f}_{kj}

表 4.11: 1001 から 1251 タイムステップにおける混合比の推定結果 \hat{a}_k

	\hat{a}_k
$k = 1$	0.07
$k = 2$	0.20
$k = 3$	0.23
$k = 4$	0.07
$k = 5$	0.23
$k = 6$	0.10
$k = 7$	0.10

表 4.12: 1001 から 1251 タイムステップにおけるクラスタの推定結果

\hat{k}	車両番号	クラスタに含まれる車両数
1	12	1
2	7 8 13 18 21	5
3	1 5 6 9 17 23	6
4	10	1
5	3 4 11 14 16 22	6
6	2 19	2
7	15 20	2

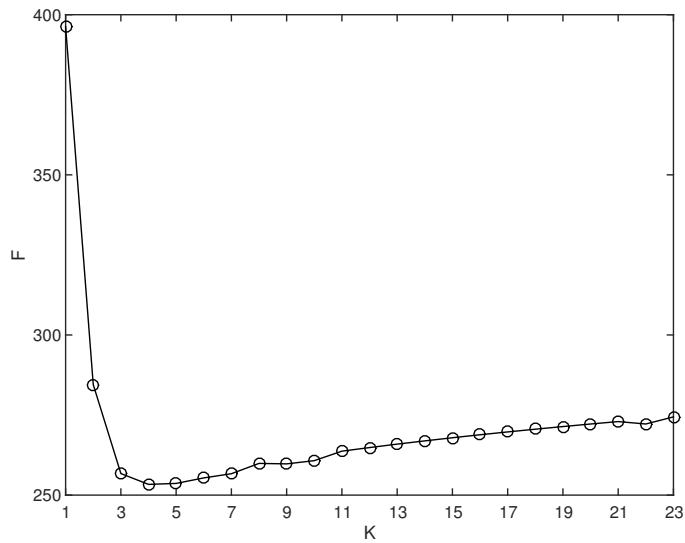


図 4.14: 1251 から 1501 タイムステップにおける混合数に対する $\mathcal{F}(p_{vb})$ の値 ($K = 4$ が最小)

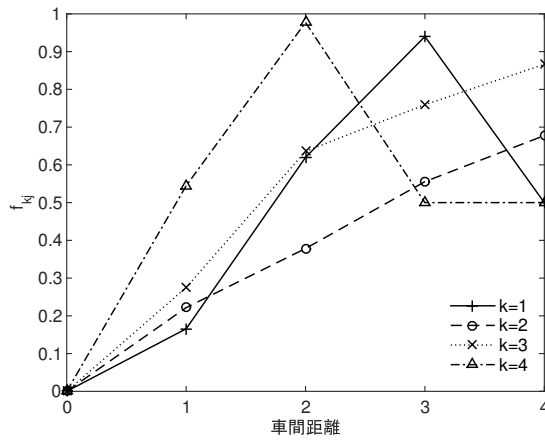


表 4.13: 1251 から 1501 タイムステップにおける混合比の推定結果 \hat{a}_k

	\hat{a}_k
$k = 1$	0.19
$k = 2$	0.19
$k = 3$	0.40
$k = 4$	0.22

図 4.15: 1251 から 1501 タイムステップにおける OV 関数の推定結果 \hat{f}_{kj}

表 4.14: 1251 から 1501 タイムステップにおけるクラスタの推定結果

\hat{k}	車両番号	クラスタに含まれる車両数
1	5 6 8 23	4
2	7 10 12 13	4
3	1 3 4 9 11 16 17 18 21 22	10
4	2 14 15 19 20	5

4.5 考察

ここでは、推定結果の確からしさに関する考察と多種粒子 TASEP, 多種粒子 ZRP の推定結果がどのように異なり、どのような場合にそれぞれの推定方法が適しているかということ考察する。

4.5.1 推定結果の確からしさ

まずパラメータ推定結果の確からしさについて考察する。前節における人工データと実測データにおける OV 関数の推定結果において、図 4.3 の $k = 4$ の車間距離 4 や図 4.5 の $k = 6$ の車間距離 4 や $k = 5$ の車間距離 4 において、減少している部分が見られる。これは推定に用いたデータ数に起因することを以下に示す。

多種粒子 ZRP の統計的表現は試行回数が可変な 2 項分布の積を混合した分布である。2 項分布において成功確率はパラメータであり、これは試行回数と成功回数から推定される。多種粒子 ZRP においても x^n と y^n によって成功確率が推定されるため、推定の確からしさはこれらの回数に影響を受ける。特に、試行回数の増加が成功回数の増加につながり、成功確率の確からしさにつながると考えられる。そこで 1 から 251 タイムステップについて推定結果の試行回数を調べる。ちなみに試行回数が 0 の場合、成功確率は事前分布のみに依存するため $\alpha/(\alpha + \beta)$ となり、今回の場合、0.5 と推定される。

i 番目の車両に k 番目の OV 関数が割り当てられる確率の推定値は \hat{z}_{ik} であるので、 k 番目の OV 関数の車間距離 j の推定に用いられた試行回数の平均は

$$\hat{x}_{kj} = \sum_{i=1}^n \hat{z}_k^{(i)} x_j^{(i)}.$$

この平均値が小さい場合、前進確率は十分なデータにより推定が行われなかったと考えられる。1 から 251 タイムステップにおける \hat{x}_{kj} を表 4.15 にまとめる。この表から OV 関数の減少部分は \hat{x}_{kj} が小さいことがわかる。

したがって、試行回数の増加により推定結果の確からしさは向上することが期待できる。試行回数を増加させるため、人工データの CA の更新回数を 100 回から

表 4.15: 1 から 251 タイムステップにおける \hat{x}_{kj} の値

\hat{x}_{kj}	$j = 1$	$j = 2$	$j = 3$	$j = 4$
$k = 1$	18.0	70.0	84.0	78.0
$k = 2$	552.6	766.4	132.0	9.0
$k = 3$	714.0	477.7	5.0	0
$k = 4$	890.1	178.0	0	0
$k = 5$	691.8	74.0	2	0
$k = 6$	193.0	0	0	0

1000 回に変化させたときに得られたパラメータの平均値の結果を図 4.16, 表 4.16 に示す. 更新回数が 100 回である場合の結果である図 4.3 と表 4.2 に比べ, $k = 4$ の車間距離 4 についても正しく推定できており, 真のパラメータに近いものになっていることがわかる.

さらに, 実測データについてもタイムステップを増やして推定を行った. 1 から 501 タイムステップにおける OV 関数の推定結果とそのときの \hat{x}_{kj} の値を示す. 図 4.17 が OV 関数の推定結果 \hat{f}_{kj} であり, 表 4.17 が混合比の推定結果 \hat{a}_k を表し, 表 4.18 が \hat{x}_{kj} の値である. なお, 1 から 251 タイムステップの場合と比較するため, ここでは $K = 6$ で固定して推定を行っている. OV 関数の減少部分が 1 から 251 タイムステップにおける推定結果である図 4.5 に比べ少なくなっていること, \hat{x}_{kj} の値が増加していることから結果の信頼性が向上していると考えられる.

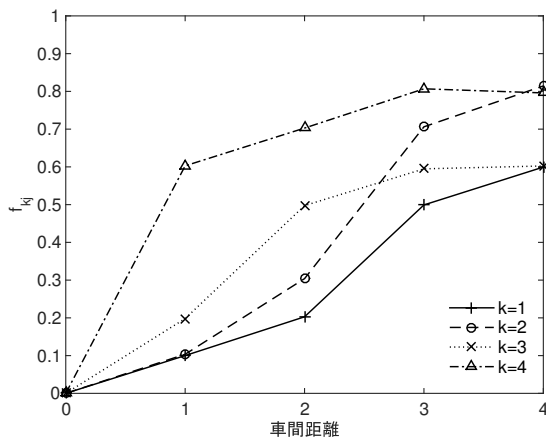


図 4.16: 人工データの更新回数を 100 回から 1000 回に変えた時の OV 関数の推定結果 \hat{f}_{kj}

表 4.16: 人工データの更新回数を 100 回から 1000 回に変えた時の混合比の推定結果 \hat{a}_k

	a_k^*
$k = 1$	0.21
$k = 2$	0.28
$k = 3$	0.25
$k = 4$	0.26

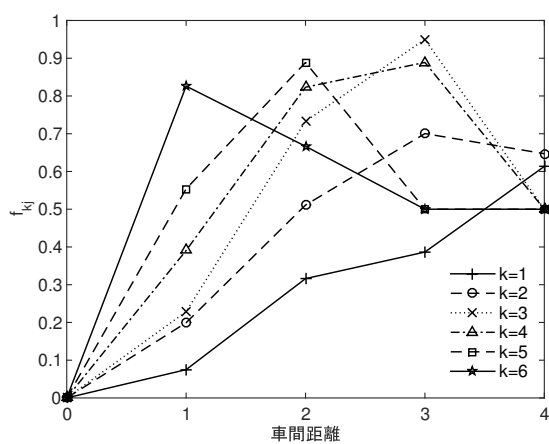


図 4.17: 1 から 501 タイムステップにおける OV 関数の推定結果 \hat{f}_{kj}

表 4.17: 1 から 501 タイムステップにおける混合比の推定結果 \hat{a}_k

	\hat{a}_k
$k = 1$	0.07
$k = 2$	0.24
$k = 3$	0.21
$k = 4$	0.24
$k = 5$	0.14
$k = 6$	0.10

表 4.18: 1 から 501 タイムステップにおける \hat{x}_{kj} の値

\hat{x}_{kj}	$j = 1$	$j = 2$	$j = 3$	$j = 4$
$k = 1$	38.0	153.0	161.0	148.0
$k = 2$	1117.0	1405.0	305.0	83.0
$k = 3$	1511.6	874.2	18.0	0
$k = 4$	2119.4	444.8	16.0	0
$k = 5$	1038.0	25.0	0	0
$k = 6$	389.0	1	0	0

4.5.2 多種粒子 TASEP と多種粒子 ZRP の比較

ここでは、1 から 251 タイムステップの推定結果を用いて多種粒子 TASEP と多種粒子 ZRP の比較を行い、その後今回提案した2つの推定方法のどちらがどのような意味で利点を持つかということを議論する。2つのモデルは違うモデルであり完全な比較が行えないため、ホップ確率 $\hat{f}_1, \dots, \hat{f}_2$ と表 4.15 の \hat{x}_{kj} がどのクラスタに対しても比較的大きい値を持つ車間距離 1 における OV 関数の値 $\hat{f}_{11}, \dots, \hat{f}_{K1}$ について各クラスタに属する車両をもとに比較する。表 4.19(a), 表 4.19(b) にホップ確率と各クラスタに属する車両の車両番号、車間距離 1 における OV 関数の値と属する車両の車両番号をそれぞれ示す。これらの表を比較すると、例えば車両番

ホップ確率	車両番号
$\hat{f}_1 = 0.43$	1,4,5,6,7,9,10,11, 12,13,16,17,18,21, 22,23
$\hat{f}_2 = 0.53$	2,3,8,15,19
$\hat{f}_3 = 0.87$	14,20

(a) 多種粒子 TASEP の各ホップ確率と属する車両番号

OV 関数	車両番号
$\hat{f}_{11} = 0.15$	7
$\hat{f}_{21} = 0.21$	4,5,10,11,13,18
$\hat{f}_{31} = 0.22$	1,6,17,21,23
$\hat{f}_{41} = 0.39$	8,9,12,16,22
$\hat{f}_{51} = 0.51$	2,3,15,19
$\hat{f}_{61} = 0.87$	14,20

(b) 多種粒子 ZRP の車間距離 1 における各 OV 関数と属する車両番号

表 4.19: 2つの推定結果における各車両に割り当てられるクラスタ

号 14, 20 の車両はホップ確率が高いクラスタと OV 関数の値が高いクラスタに属しており、このような対応関係が車両番号 2, 3, 15, 19 などにおいても確認することができるため、多種粒子 ZRP の推定結果が多種粒子 TASEP の推定結果と整合性のあるものであることが確認できる。さらに、ホップ確率が低いクラスタである \hat{f}_1 に属する車両は OV 関数では属するクラスタが細分化されており、多種粒子 ZRP モデルにより細かいクラスタリングが可能であることがわかる。

次に、2つの推定方法のどちらがどのような意味で利点を持つか考察する。多種

粒子 ZRP は車両の位置を十分多く計測可能である場合、OV 関数の各車間距離のデータが十分にあるため精度良く推定可能であると考えられる。逆に、計測が十分に行えない場合、データが車間距離で分散してしまいどの車間距離においても推定が十分に行えず精度が悪くなると考えられる。一方、多種粒子 TASEP は車両の位置を十分多く計測可能である場合、車間距離について平均化を行ってしまうため精度が悪くなると考えられるが、計測が十分に行えない場合、データが車間距離で分散しないため多種粒子 ZRP に比べ精度良く推定可能であると考えられる。したがって、計測範囲が狭い場合の推定は多種粒子 TASEP が適しており、広い場合の推定は多種粒子 ZRP が適していると推察される。

第5章 他の学習法との比較

ここまでで、多種粒子 TASEP と多種粒子 ZRP モデルに変分ベイズ法を適用する方法について考えた。変分ベイズ法などの学習法により得られる推定結果を用いて、推定で用いられなかった新しい車両の動きを高精度で少ない演算量で予測することは、CA モデルにより交通流のシミュレーションを行う上で重要である。そこで本章では、変分ベイズ法、最尤法、ベイズ法の比較を予測精度と計算時間の観点で行う。

今回は真の分布が多種粒子 ZRP である場合を考え、多種粒子 ZRP の学習モデル $p(y|x, w)$ を用いて学習を行い真の分布を予測する。予測精度として、真のパラメータ w^* と車両密度 ρ^* が与えられたときの真の分布 $q(y|x, \rho^*)q(x|\rho^*) = p(y|x, \rho^*, w^*)q(x|\rho^*)$ を用いて、以下の汎化誤差 $G_n(x^n, y^n)$ を計算する。

$$G_n(x^n, y^n) = \int p(y|x, \rho^*, w^*)q(x|\rho^*) \log \frac{p(y|x, \rho^*, w^*)}{\hat{p}(y|x, \rho^*, x^n, y^n)} dx dy.$$

汎化誤差の数値計算を行うため、ここではテストデータ $(x^T, y^T) = (x^{(t)}, y^{(t)}; t = 1, \dots, T)$ を真の分布から発生させ、次の $G'_n(x^n, y^n)$ を計算する。

$$G'_n(x^n, y^n) = \frac{1}{T} \sum_{t=1}^T \log \frac{p(y^{(t)}|x^{(t)}, \rho^*, w^*)}{\hat{p}(y^{(t)}|x^{(t)}, \rho^*, x^n, y^n)} \quad (5.1)$$

ここで、 $\hat{p}(y|x, \rho^*, x^n, y^n)$ は各推定手法によって得られる予測分布であり、変分ベイズ法は、式 (4.16) によって得られる分布である。最尤法により得られる分布は、EM 法を用いることで得られるパラメータ \hat{w}_{em} を用いて $p(y|x, \hat{w}_{em})$ である。ベイズ法により得られる分布は、式 (3.9) であり、MCMC の手法 1 つであるギブスサンプリングで与えられる事後分布のサンプルを $(w^l; l = 1, \dots, L)$ として

$$p'_{bayes}(y|x, x^n, y^n) = \frac{1}{L} \sum_{l=1}^L p(y|x, w^l) \quad (5.2)$$

を用いて予測分布とする．次に，簡単にEM法とギブスサンプリングの学習アルゴリズムについて説明する．

5.1 EM法の学習アルゴリズム

EM法は最尤推定量の局所最適解を求める方法であり，今回の場合隠れ変数の推定とパラメータの推定を交互に行うアルゴリズムである．今回は， z^n の初期値 $\hat{z}^n = (\hat{z}_k^{(i)}; i = 1, \dots, n, k = 1, \dots, K)$ を定め，パラメータを

$$\hat{w} = \arg \max_w \log p(y^n, \hat{z}^n | x^n, w)$$

により更新する． $\hat{w} = (\hat{a}_k, \hat{f}_{kj}; k = 1, \dots, K, j = 1, \dots, M)$ を用いて

$$p(z^n | x^n, y^n, \hat{w}) = \frac{p(z^n, y^n | x^n, \hat{w})}{p(y^n | x^n, \hat{w})} = \prod_{i=1}^n \frac{p(z^{(i)}, y^{(i)} | x^{(i)}, \hat{w})}{p(y^{(i)} | x^{(i)}, \hat{w})},$$

$$\hat{z}_k^{(i)} = p(z_k^{(i)} = 1, y^{(i)} | x^{(i)}, \hat{w}) = \frac{p(z_k^{(i)} = 1, y^{(i)} | x^{(i)}, \hat{w})}{p(y^{(i)} | x^{(i)}, \hat{w})}$$

により $\hat{z}_k^{(i)}$ を更新し，これらを繰り返すことで \hat{w}_{em} を求める．多種粒子ZRPの場合，具体的なアルゴリズムは以下ようになる．

アルゴリズム 3 (多種粒子ZRPのEM法のアルゴリズム)．

1. $\{\hat{z}_k^{(i)}; i = 1, \dots, n, k = 1, \dots, K\}$ を初期化する．

2. M-ステップ: \hat{w} の更新

$$\hat{a}_k \leftarrow \frac{\sum_{i=1}^n \hat{z}_k^{(i)}}{n},$$

$$\hat{f}_{kj} \leftarrow \frac{\sum_{i=1}^n y_j^{(i)} \hat{z}_k^{(i)}}{\sum_{i=1}^n x_j^{(i)} \hat{z}_k^{(i)}}. \quad (5.3)$$

3. E-ステップ: $p(z^n | x^n, y^n, \hat{w})$ の更新

$$L_{ik} \leftarrow \log \hat{a}_k + \sum_{j=1}^M \left\{ y_j^{(i)} \log \hat{f}_{kj} + (x_j^{(i)} - y_j^{(i)}) \log(1 - \hat{f}_{kj}) \right\}, \quad (5.4)$$

$$\hat{z}_k^{(i)} \leftarrow \frac{\exp(L_{ik})}{\sum_{l=1}^K \exp(L_{il})}.$$

4. E-ステップと M-ステップを交互に繰り返し更新する.

ここで、式 (5.3) は分母、分子の項に $\hat{z}_k^{(i)}$ を含んでおり、推定途中で分母、分子が 0 になる場合がある。そのため、本論文では分母と分子が 0 になった場合、0 になる 1 つ前の結果を用いて推定結果とする。また、得られた推定結果は初期値に依存して変わるため、各初期値で得られた \hat{w} に対して式 (5.4) を用いて

$$p(x^n | \hat{w}) = \sum_{i=1}^n \log \sum_{k=1}^K \exp(L_{ik}) \quad (5.5)$$

の値が最も大きいものを最終的な推定結果として用いた。

5.2 ギブスサンプリングの学習アルゴリズム

ギブスサンプリングは今回の場合、パラメータと隠れ変数の同時事後分布 $p(w, z^n | x^n, y^n)$ からのサンプリングを行うため、 $p(z^n | w, x^n, y^n)$ からのサンプリングと $p(w | z^n, x^n, y^n)$ からのサンプリングを交互に行う手法である。ここで、 z^n と w のサンプリングの順序は確率的に決定しなければ詳細つり合い条件が満たされないが、式 (5.2) は $L \rightarrow \infty$ で式 (3.9) に収束するため、今回は同じ順番でサンプリングを行う。

サンプリングを行うため、 $p(w | z^n, x^n, y^n)$ と $p(z^n | w, x^n, y^n)$ の確率分布について考える。事前分布として変分ベイズ法の場合と同じ式 (4.5) を用いると、 $p(w | z^n, x^n, y^n)$

は

$$\begin{aligned}
p(w|z^n, x^n, y^n) &\propto p(y^n, z^n|w, x^n)\varphi(w) \\
&\propto \prod_{k=1}^K a_k^{\sum_{i=1}^n z_k^{(i)} + \phi - 1} \prod_{j=1}^M f_{kj}^{\sum_{i=1}^n z_k^{(i)} y_j^{(i)} + \alpha - 1} (1 - f_{kj})^{\sum_{i=1}^n z_k^{(i)} (x_j^{(i)} - y_j^{(i)}) + \beta - 1} \\
&\propto \text{Dir}(a|\{\sum_{i=1}^n z_k^{(i)} + \phi\}_{k=1}^K) \\
&\quad \prod_{k=1}^K \prod_{j=1}^M \text{Beta}(f_{kj}|\sum_{i=1}^n z_k^{(i)} y_j^{(i)} + \alpha, \sum_{i=1}^n z_k^{(i)} (x_j^{(i)} - y_j^{(i)}) + \beta)
\end{aligned}$$

となるため、 $p(w|z^n, x^n, y^n)$ からのサンプリングはディリクレ分布とベータ分布を用いることで可能である。また $p(z^n|w, x^n, y^n)$ は

$$\begin{aligned}
p(z^n|w, x^n, y^n) &\propto p(y^n, z^n|w, x^n) \\
&\propto \prod_{i=1}^n \prod_{k=1}^K \exp(z_k^{(i)} L_{ik}), \\
L_{ik} &= \log a_k + \sum_{j=1}^M \left\{ y_j^{(i)} \log f_{kj} + (x_j^{(i)} - y_j^{(i)}) \log(1 - f_{kj}) \right\}
\end{aligned}$$

となり、 $z^{(i)} \in \mathcal{C}$ が $z_k^{(i)} = 1$ である確率は

$$p(z_k^{(i)} = 1|w, x^n, y^n) = \frac{\exp(L_{ik})}{\sum_{k=1}^K \exp(L_{ik})}$$

であるため、この確率に従って z^n のサンプリングが可能である。以上をまとめると、アルゴリズムは次のようになる。

アルゴリズム 4 (多種粒子 ZRP のギブスサンプリングのアルゴリズム).

1. $\{z_k^{(i)}; i = 1, \dots, n, k = 1, \dots, K\}$ を初期化する。

2. $p(w|z^n, x^n, y^n)$ からのサンプリング:

以下の確率分布に従って混合比 $a = (a_1, \dots, a_K)$ と OV関数 f_{kj} をサンプリングする。

$$a \sim \text{Dir}(a | \{\sum_{i=1}^n z_k^{(i)} + \phi\}_{k=1}^K),$$

$$f_{kj} \sim \text{Beta}(f_{kj} | \sum_{i=1}^n z_k^{(i)} y_j^{(i)} + \alpha, \sum_{i=1}^n z_k^{(i)} (x_j^{(i)} - y_j^{(i)}) + \beta).$$

3. $p(z^n | w, x^n, y^n)$ からのサンプリング:

以下の L_{ik} を用いて隠れ変数 z^n をサンプリングする.

$$L_{ik} \leftarrow \log a_k + \sum_{j=1}^M \left\{ y_j^{(i)} \log f_{kj} + (x_j^{(i)} - y_j^{(i)}) \log(1 - f_{kj}) \right\},$$

$$z^n \sim \prod_{i=1}^n \prod_{k=1}^K \frac{\exp(z_k^{(i)} L_{ik})}{\sum_{l=1}^K \exp(L_{il})}.$$

4. 2. と 3. を交互に繰り返す.

なお、ギブスサンプリングは初期値の依存性を無くするため、初期のサンプリングを捨てる期間であるバーンイン、サンプリング間の相関を無くするためのサンプリング間隔を設定する必要がある。

5.3 実験設定と実験結果

ここまでで各手法における予測分布の計算方法を説明したので、汎化誤差の計算を実験的に行う。今回学習データを生成する真のパラメータ w^* として表5.1のものを用いた。この真のパラメータを用いて100台 ($n = 100$) の車両を車両密度が0.5の

表 5.1: 汎化誤差の計算における真のパラメータ

	a_k^*	f_{k1}^*	f_{k2}^*	f_{k3}^*
$k = 1$	0.5	0.2	0.4	0.6
$k = 2$	0.5	0.5	0.7	0.9

CA 上を100回更新することで学習データ (x^n, y^n) を生成した。学習データに対し

て、変分ベイズ法は1000回の繰り返しを100個の初期値に対して行い、式(4.17)を最小にするものを学習結果として用いた。ハイパーパラメータ $(\phi, \alpha, \beta) = (1, 1, 1)$ とした。最尤法は1000回の繰り返しを100個の初期値に対して行い、式(5.5)を最大にするものを推定結果として用いた。また、OV関数の値が不定になった場合、1つ前の学習結果を用いて推定を行った。ベイズ法はギブスサンプリングのバーンインを1000回、サンプリングする間隔を200回ごととして事後分布からのサンプリングを合計1000個取得した。また、ハイパーパラメータは変分ベイズ法と同じものを使用し $(\phi, \alpha, \beta) = (1, 1, 1)$ とした。混合数 K は手法ごとに変わらず、真のパラメータと同じ $K = 2$ の場合と冗長である $K = 4$ の場合を用いて推定を行った。テストデータは、10000台 ($T = 10000$) の車両を車両密度が0.5のCA上を100回更新することで生成し、それぞれの手法で汎化誤差を式(5.1)により計算した。また、汎化誤差は学習データ、テストデータによって値がゆらぐため、学習データ、テストデータを100セット用意しそれぞれで汎化誤差の計算を行い、平均値とヒストグラムの比較をそれぞれの手法で行った。

この実験設定の基で、まず真のパラメータに対して冗長でない $K = 2$ の場合の各手法の汎化誤差を図5.1に示す。(a)が変分ベイズ法、(b)が最尤法、(c)がベイズ法のデータセット毎のヒストグラムを表しており、横軸が汎化誤差の値、縦軸が頻度である。各ヒストグラム上部に表記した $E[G]$ の値が汎化誤差のデータセット平均の値を示している。この結果から、最尤法は他の2つの推定手法に比べデータセットによって汎化誤差の精度が悪くなる場合があり、それが原因で汎化誤差の平均値が大きくなっていることがわかる。また、変分ベイズ法とベイズ法はほぼ同等の平均汎化誤差を持つことが分かる。

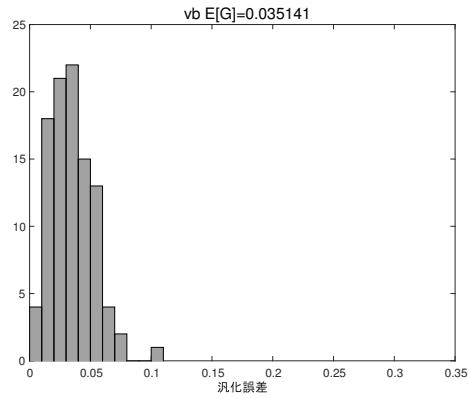
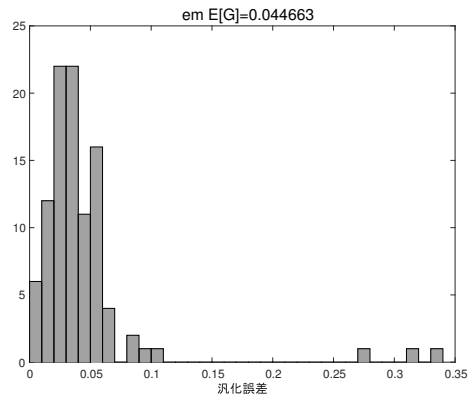
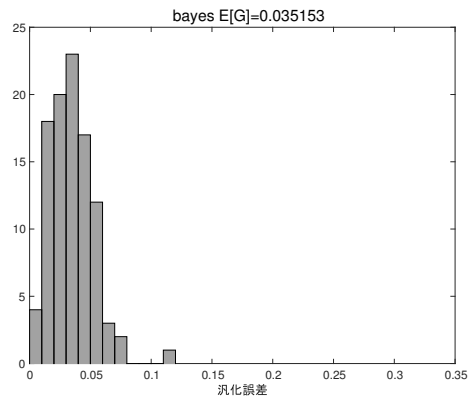
次に、図5.2に $K = 2$ で各データセットに対して各推定手法で要した計算時間のヒストグラムを示す。(a)が変分ベイズ法、(b)が最尤法、(c)がベイズ法の計算時間を示しており、横軸が計算時間、縦軸が頻度を表し図の上部にそれぞれの推定で要した時間の平均値を表記している。この結果から、ベイズ法は他の2つの手法に比べ大きく時間を要していることが分かる。また、最尤法は0秒付近で推定が終了するデータセットを確認することができるが、これはそのデータセット

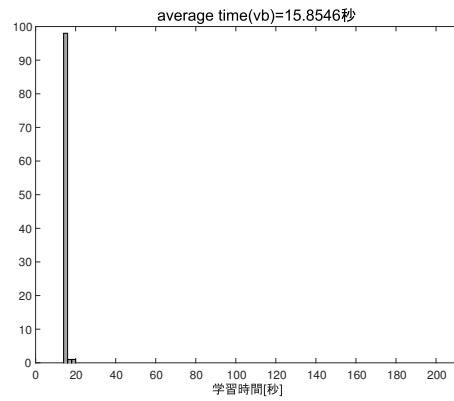
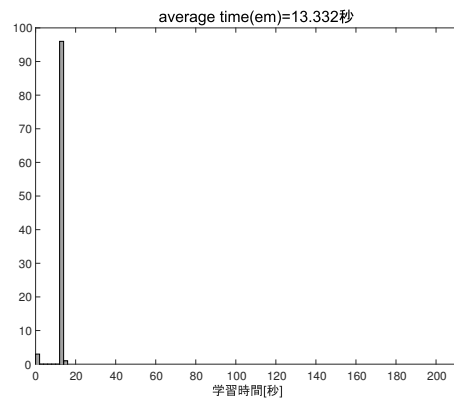
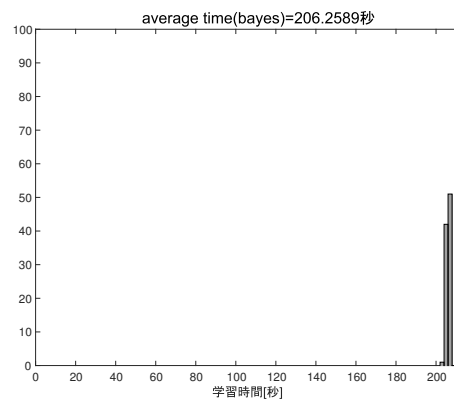
に対して初期値を変えても繰り返し途中で OV 関数が不定になってしまい，途中で推定が打ち切られたことが原因であると考えられる．

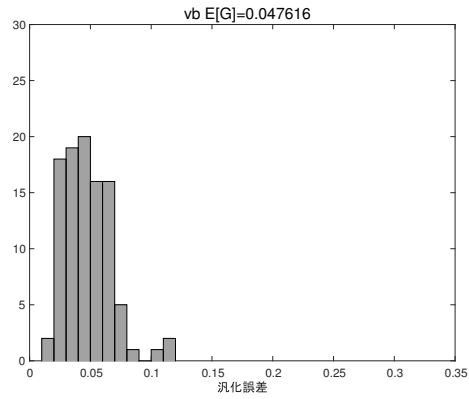
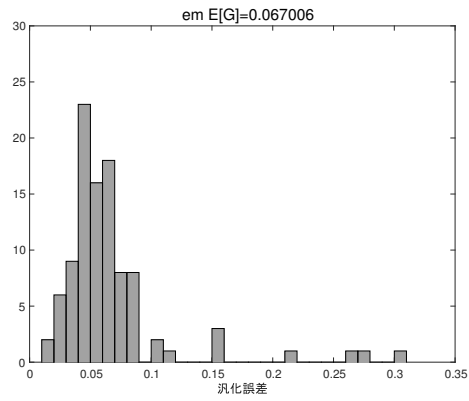
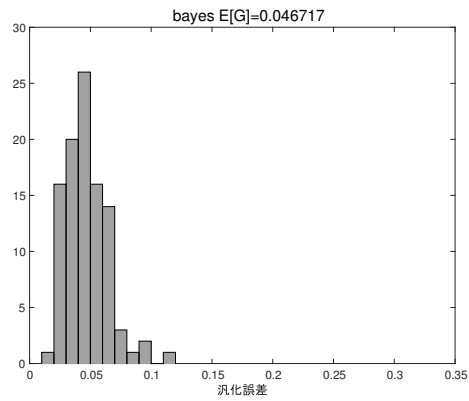
同様にして真のパラメータに対して冗長である $K = 4$ における汎化誤差のヒストグラムを図 5.3 に示す．この場合においても，変分ベイズ法はベイズ法とほぼ同じ予測精度を持つことがわかった．また，最尤法は汎化誤差が大きくなるデータセットがさらに増えてしまい，他の手法に比べ精度が悪いことがわかった．

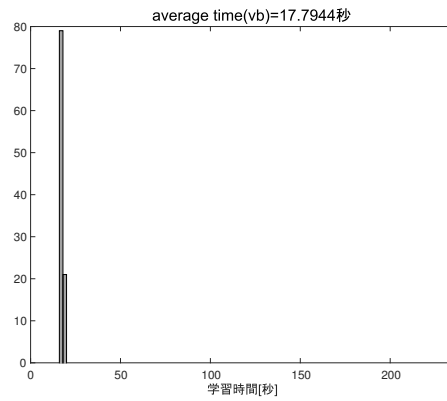
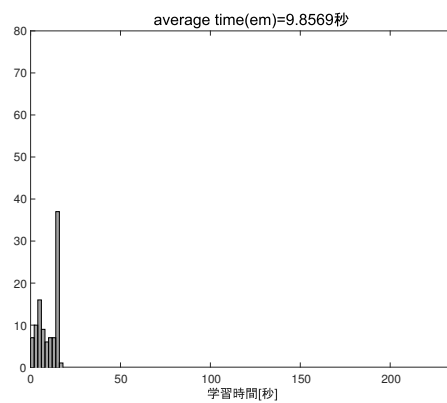
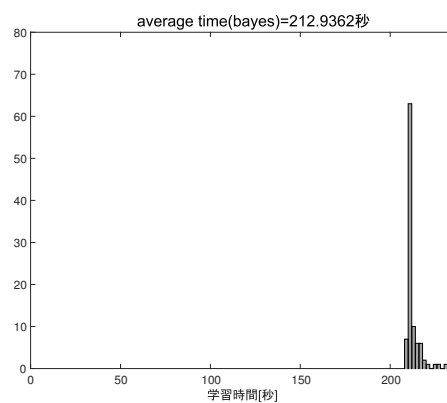
$K = 4$ における計算時間は図 5.4 であり，この場合においてもベイズ法は他の 2 つの手法に比べ大きく時間が掛かっていることがわかる．また，最尤法は計算時間が 0 秒付近である頻度が図 5.2 の場合よりも増えており，繰り返し途中で推定が終わる頻度が増えることが分かった．

以上の結果から，ベイズ法と変分ベイズ法は最尤法に比べ予測精度が高く，2 つは同等の予測精度を持ち，変分ベイズ法と最尤法はベイズ法に比べ計算時間が少なく，2 つは同等の計算時間を持つことが実験的に分かった．以上から，予測精度・演算量の観点で変分ベイズ法が他の推定手法に比べ有用であることが分かった．

(a) $K = 2$ における変分ベイズ法の汎化誤差のヒストグラム(b) $K = 2$ における最尤法の汎化誤差のヒストグラム(c) $K = 2$ におけるベイズ法の汎化誤差のヒストグラム図 5.1: $K = 2$ における各推定手法における汎化誤差のヒストグラム

(a) $K = 2$ における変分ベイズ法の計算時間のヒストグラム(b) $K = 2$ における最尤法の計算時間のヒストグラム(c) $K = 2$ におけるベイズ法の計算時間のヒストグラム図 5.2: $K = 2$ における各推定手法の計算時間のヒストグラム

(a) $K = 4$ における変分ベイズ法の汎化誤差のヒストグラム(b) $K = 4$ における最尤法の汎化誤差のヒストグラム(c) $K = 4$ におけるベイズ法の汎化誤差のヒストグラム図 5.3: $K = 4$ における各推定手法における汎化誤差のヒストグラム

(a) $K = 4$ における変分ベイズ法の計算時間のヒストグラム(b) $K = 4$ における最尤法の計算時間のヒストグラム(c) $K = 4$ におけるベイズ法の計算時間のヒストグラム図 5.4: $K = 4$ における各推定手法の計算時間のヒストグラム

第6章 結論

本論文では、SCA モデルの一種である多種粒子 TASEP と多種粒子 ZRP の統計的推測に変分ベイズ法を適用することを提案し、その学習アルゴリズムの導出を行った。また、提案法を人工データ、実測データに適用し他の学習法と比較することで有効性の検証を行った。

学習アルゴリズムの導出に関して、具体的なアルゴリズムの構成及び推定結果の評価を行う評価関数を具体的に示すことができたため、提案法により、実際に多種粒子 TASEP と多種粒子 ZRP モデルの推定を変分ベイズ法で行う上で必要な手続きを全て記述することができた。有効性の検証に関して、人工データだけでなく実測データに適用を行い、データに対して適切な推定が行われていること、運転方法の多様性や非定常性を捉えることができた。また、数値実験によって汎化誤差と計算時間を学習法ごとに測ることで実験的に変分ベイズ法の予測性能と演算量における有用性を検証することが可能となった。

本論文は、1章では、交通流解析に関する社会的背景および学術的背景について述べ、本研究の目的を明らかにした。2章では、3,4章で検討したモデルの位置づけを確認するため、様々な CA モデルとその関係性について述べた。3章では、多種粒子 TASEP モデルのパラメータを変分ベイズ法で推定するため学習アルゴリズムの導出を行い、人工データ、実測データに対して適用した。4章では、多種粒子 ZRP モデルのパラメータを変分ベイズ法で推定するため学習アルゴリズムの導出を行い、人工データ、実測データに対して適用した。5章では、最尤法、ベイズ法と変分ベイズ法の比較を多種粒子 ZRP モデルを用いて行い、予測精度と演算量の観点で変分ベイズ法が有用であるということを実験的に示した。

交通流モデルの統計的推測は詳細な交通流解析を行う上で重要であり、交通流

分野の研究者に対して徐々に認識されている [88]. 特に, 本論文で考えた SCA モデルは基本的な交通流モデルであり, 本論文は交通流モデルに統計的推測を適用するための橋頭保を築いただけでなく, 交通流分野における解析方法の視野を広げる貢献となった.

今後の展望は交通流モデル, 実用化, 渋滞検出の3つの観点が考えられる. 交通流モデルの観点では, (1)ZRP の単調性を保証するモデルの推定, (2)今回考えたような多様性を含む SCA モデルを複雑にしたモデルへ統計的推測を適用することが挙げられる. (1)について, 本論文では K 個の OV 関数 f_{kj} を車間距離 j に独立な形のを定義していたため, 統計的表現がパラメータの積の形となり, 変分ベイズ法のアルゴリズムが導出できた. これに対し OV 関数の単調増加性を保証するモデルとして例えば, 式 (2.1) や $f_{kj} = \sum_{l=1}^j \alpha_{kl}$ で α_{kl} が正値をとるパラメータなどが考えられる. これらのモデルは, 本論文で考えた変分ベイズ法のアルゴリズムを同じように適用することができないが, 変分ベイズ法が適用可能な近似を考案するか, Automatic Differentiation Variational Inference(ADVI) [44] といったアルゴリズムの導出を行わず変分ベイズ法の推測を行う方法を用いるか, 変分ベイズ法以外の学習法により推定は可能であると考えられるため, それらの学習法を用いてパラメータを推定することが今後の課題である. (2)について, 例えば TASEP や ZRP を発展させたモデルに Stochastic Optimal Velocity(SOV) と呼ばれるモデルがあり [35], そのモデルで多様性を考慮したモデルである多種粒子 SOV モデルが考えられる. SOV モデルは時刻 t における車間距離が x_t のとき前進確率 $p_t(x_t)$ が以下のように変化する.

$$p_t(x_t) = (1 - a)p_{t-1} + aV(x_t)$$

ここで, $V(\cdot)$ は OV 関数であり, $0 \leq a \leq 1$ である. このモデルは $a = 0$ のとき ZRP と一致し, $a = 1$ のとき TASEP と一致するモデルであり, 時刻 $t = 0$ における前進確率 p_0 , $V(\cdot)$, a が推定するパラメータである. このモデルに従い CA 上を走行する車両は TASEP, ZRP の場合と同じくベルヌーイ試行に従った振る舞いを行っており, 時刻 t における前進を $y_t = 1$, 同じセルに留まることを $y_t = 0$ で

表すと SOV の統計的表現は

$$p(y_1, \dots, y_T | x_1, \dots, x_T) = \prod_{t=1}^T p_t(x_t)^{y_t} (1 - p_t(x_t))^{1-y_t}$$

で表すことができる。多種粒子 SOV はこのモデルの混合により表現することができる。このモデルも前述の単調性を仮定した ZRP と同じく今回考えた変分ベイズ法のアルゴリズムを適用することができないため、新しく学習法を考えることが今後の課題である。多種粒子 SOV モデル以外に、NS モデルや時間とともに運転方法が変化するモデル、2車線以上で車線変更が発生するモデルなどが考えられる。こういったモデルについても各車両の動きをベルヌーイ試行で記述することができ、統計的表現を構築することができると考えられるため、今後発展することを期待したい。

次に、今回の提案法を実用化する場合の課題については、(1) 学習データを得ること、(2) SCA モデルの離散幅が課題になると考えられる。(1) に関して、例えば、高速道路を撮影した動画から学習データを得ることを考えると、画像中の車両の特定とフレーム間での車両の同定が必要である。一方、実際の道路では車両の追い越しや障害物、霧や雨などの天候により車両の同定が難しくなる場合があるため、そのような場合の学習データの抽出が今後の課題である。(2) に関して、本論文は、1時刻で2つ以上のセルの移動を行うといったモデルの動きに適さないものを例外として扱い例外が少ない離散幅を用いて推定を行った。一方、このような例外は実データを扱う上でたびたび発生するものであるため、例外を含むことが可能な SCA モデルを考えることが再現性の高いモデルを構築する上で今後の課題である。さらに、そのようなモデルでは離散幅により推定精度の高いモデル構築が可能となると考えられるため、推定精度に関する離散幅の最適化も今後発展することを期待したい。

最後に、統計的推測によって渋滞検出する方法について考えると、渋滞の発生前後で運転方法は変わると考えられるため、例えば混合数 K の変化や仮説・検定により運転方法が変化したかどうか調べることができるはずであり、そのような方法で渋滞検出を行うことが今後の課題である。

謝辞

本研究では Suigyama らによって行われた実験で得られた実測データを用いた [71]. ここに謝意を表します.

本研究の遂行ならびに論文の作成にあたり, 懇切なる御指導を賜りました産業技術総合研究所主任研究員 山崎啓介博士、東京工業大学大学院総合理工学研究科教授 渡辺澄夫先生に謹んで感謝の意を表します。

またこれまで研究を通して活発な議論にお付き合い頂いた渡辺研究室の山田耕史氏、幸島匡宏氏、山田哲太郎氏、三木拓史氏、小林浩一氏、加藤尊氏、大工廻和美氏、大原成裕氏、玉井雄介氏、司空舜氏、宮崎大氏、香田夏輝氏、近藤稔彦氏、須貝将士氏、竹内理人氏、林直輝氏、岡野高志氏、藤村光氏の皆様に感謝致します。

2017年3月 中村 文士

関連図書

- [1] M. S. Alber, M. A. Kiskowski, J A. Glazier, and Y. Jiang. On Cellular Automaton Approaches to Modeling Biological Cells. pp. 1–39. Springer, 2003.
- [2] G. B. Arous and I. Corwin. Current fluctuations for TASEP: A proof of the Prähofer-Spohn conjecture. *The Annals of Probability*, Vol. 39, No. 1, pp. 104–138, jan 2011.
- [3] H. Attias. *Inferring parameters and structure of latent variable models by variational Bayes*. Morgan Kaufmann Publishers, 1999.
- [4] A. Ayyer and S. Linusson. An inhomogeneous multispecies TASEP on a ring. *Advances in Applied Mathematics*, Vol. 57, pp. 21–43, 2014.
- [5] M. Bando, K. Hasebe, A. Nakayama, A. Shibata, and Y. Sugiyama. Dynamical model of traffic congestion and numerical simulation. *Physical Review E*, Vol. 51, No. 2, pp. 1035–1042, feb 1995.
- [6] R. Barlovic, L. Santen, A. Schadschneider, and M. Schreckenberg. Metastable states in cellular automata for traffic flow. *The European Physical Journal B*, Vol. 5, No. 3, pp. 793–800, oct 1998.
- [7] M. J. Beal. *Variational algorithms for approximate Bayesian inference*. 2003.
- [8] O. Biham, A. A. Middleton, and D. Levine. Self-organization and a dynamical transition in traffic-flow models. *Physical Review A*, Vol. 46, No. 10, pp. R6124–R6127, nov 1992.

- [9] C. Burstedde, K. Klauck, A. Schadschneider, and J. Zittartz. Simulation of pedestrian dynamics using a two-dimensional cellular automaton. *Physica A: Statistical Mechanics and its Applications*, Vol. 295, No. 3, pp. 507–525, 2001.
- [10] C. Chen, J. Li, and D. Zhang. Study on evacuation behaviors at a T-shaped intersection by a force-driving cellular automata model. *Physica A: Statistical Mechanics and its Applications*, Vol. 391, No. 7, pp. 2408–2420, 2012.
- [11] D. Chowdhury, L. Santen, and A. Schadschneider. Statistical physics of vehicular traffic and some related systems. *Physics Reports*, Vol. 329, No. 4, pp. 199–329, 2000.
- [12] D. Chowdhury, D. E. Wolf, and M. Schreckenberg. Particle hopping models for two-lane traffic with two kinds of vehicles: Effects of lane-changing rules. *Physica A: Statistical Mechanics and its Applications*, Vol. 235, No. 3, pp. 417–439, 1997.
- [13] A. Corduneanu and C. Bishop. Variational Bayesian Model Selection for Mixture Distributions. *Artificial intelligence and Statistics*, pp. 27–34, 2001.
- [14] M. Cowles and B. P. Carlin. Markov Chain Monte Carlo Convergence Diagnostics: A Comparative Review. *Journal of the American Statistical Association*, Vol. 91, No. 434, pp. 883–904, jun 1996.
- [15] B. Derrida, M. R. Evans, V. Hakim, and V. Pasquier. Exact solution of a 1D asymmetric exclusion model using a matrix formulation. *Journal of Physics A: Mathematical and General*, Vol. 26, No. 7, pp. 1493–1517, apr 1993.
- [16] M. R. Evans. Phase transitions in one-dimensional nonequilibrium systems. *Brazilian Journal of Physics*, Vol. 30, No. 1, pp. 42–57, mar 2000.

- [17] M. R. Evans and T. Hanney. Phase transition in two species zero-range process. *Journal of Physics A: Mathematical and General*, Vol. 36, No. 28, pp. L441–L447, jul 2003.
- [18] M. R. Evans and T. Hanney. Nonequilibrium statistical mechanics of the zero-range process and related models. *Journal of Physics A: Mathematical and General*, Vol. 38, No. 19, 2005.
- [19] P. A. Ferrari, C. Kipnis, and E. Saada. Microscopic Structure of Travelling Waves in the Asymmetric Simple Exclusion Process. *The Annals of Probability*, Vol. 19, No. 1, pp. 226–244, jan 1991.
- [20] P. L. Ferrari and H. Spohn. Scaling Limit for the Space-Time Covariance of the Stationary Totally Asymmetric Simple Exclusion Process. *Communications in Mathematical Physics*, Vol. 265, No. 1, pp. 1–44, jul 2006.
- [21] M. Fukui and Y. Ishibashi. Traffic flow in 1D cellular automaton model including cars moving with high speed. *Journal of the Physical Society of Japan*, Vol. 65, No. 6, pp. 1868–1870, 1996.
- [22] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin. *Bayesian data analysis*. Chapman & Hall/CRC Boca Raton, FL, USA, 2014.
- [23] Z. Ghahramani and M J. Beal. Variational Inference for Bayesian Mixtures of Factor Analysers. *NIPS*, Vol. 12, pp. 449–455, 1999.
- [24] B.D. Greenshields, W. Channing, and H. Miller. A study of traffic capacity. *Highway Research Board proceedings*, Vol. 1935, , 1935.
- [25] S. Großkinsky and H. Spohn. Stationary measures and hydrodynamics of zero range processes with several species of particles. *Bulletin of the Brazilian Mathematical Society*, Vol. 34, No. 3, pp. 489–507, nov 2003.

- [26] F. L. Hall, B. L. Allen, and M. A. Gunter. Empirical analysis of freeway flow-density relationships. *Transportation Research Part A: General*, Vol. 20, No. 3, pp. 197–210, 1986.
- [27] J. A. Hartigan. A failure of likelihood asymptotics for normal mixtures. *Proceedings of the Berkeley conference in honor of Jerzy Neyman and Jack Kiefer*, Vol. 2, pp. 807—810, 1995.
- [28] D. Helbing. Gas-kinetic derivation of Navier-Stokes-like traffic equations. *Physical Review E*, Vol. 53, No. 3, pp. 2366–2381, mar 1996.
- [29] D. Helbing. Traffic and related self-driven many-particle systems. *Reviews of Modern Physics*, Vol. 73, No. 4, pp. 1067–1141, dec 2001.
- [30] M. Hoffman, F R. Bach, and D M. Blei. Online Learning for Latent Dirichlet Allocation, 2010.
- [31] T. Hosino, K. Watanabe, and S. Watanabe. Stochastic complexity of variational Bayesian hidden Markov models. In *Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005.*, Vol. 2, pp. 1114–1119. IEEE, 2005.
- [32] T. Hosino, K. Watanabe, and S. Watanabe. Free Energy of Stochastic Context Free Grammar on Variational Bayes. pp. 407–416. Springer Berlin Heidelberg, 2006.
- [33] D. Kaji, K. Watanabe, and S. Watanabe. Phase Transition of Variational Bayes Learning in Bernoulli Mixture. *Australian Journal of Intelligent Information Processing Systems*, Vol. 11, No. 4, 2010.
- [34] M. Kanai. Exact solution of the zero-range process: fundamental diagram of the corresponding exclusion process. *Journal of Physics A: Mathematical and Theoretical*, Vol. 40, No. 26, pp. 7127–7138, jun 2007.

- [35] M. Kanai, K. Nishinari, and T. Tokihiro. Stochastic optimal velocity model and its long-lived metastability. *Physical Review E*, Vol. 72, No. 3, p. 035102, sep 2005.
- [36] M. Kanai, K. Nishinari, and T. Tokihiro. Exact solution and asymptotic behaviour of the asymmetric simple exclusion process on a ring. *Journal of Physics A: Mathematical and General*, Vol. 39, No. 29, pp. 9071–9079, jul 2006.
- [37] Y. Kang and D. Sun. Lattice hydrodynamic traffic flow model with explicit drivers' physical delay. *Nonlinear Dynamics*, Vol. 71, No. 3, pp. 531–537, feb 2013.
- [38] V. Karimipour. Multispecies asymmetric simple exclusion process and its relation to traffic flow. *Physical Review E*, Vol. 59, No. 1, pp. 205–212, jan 1999.
- [39] J. Kaupužs, R. Mahnke, and R. J. Harris. Zero-range model of traffic flow. *Physical Review E*, Vol. 72, No. 5, p. 056125, nov 2005.
- [40] B. S. Kerner. Experimental Features of Self-Organization in Traffic Flow. *Physical Review Letters*, Vol. 81, No. 17, pp. 3797–3800, oct 1998.
- [41] B. S. Kerner and P. Konhäuser. Cluster effect in initially homogeneous traffic flow. *Physical Review E*, Vol. 48, No. 4, pp. R2335–R2338, oct 1993.
- [42] A. Kesting, M. Treiber, and D. Helbing. Enhanced intelligent driver model to access the impact of driving strategies on traffic capacity. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, Vol. 368, No. 1928, 2010.
- [43] K. Kobayashi and K. Yamazaki. Parameter estimation accuracy and active learning in the zero-range process. In *2012 12th International Conference on*

- Intelligent Systems Design and Applications (ISDA)*, pp. 811–814. IEEE, nov 2012.
- [44] A. Kucukelbir, R. Ranganath, A. Gelman, and D. Blei. Automatic Variational Inference in Stan, 2015.
- [45] I. Kusch and M. Markus. Mollusc Shell Pigmentation: Cellular Automaton Simulations and Evidence for Undecidability. *Journal of Theoretical Biology*, Vol. 178, No. 3, pp. 333–340, 1996.
- [46] M. J. Lighthill and G. B. Whitham. On Kinematic Waves. II. A Theory of Traffic Flow on Long Crowded Roads. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, Vol. 229, No. 1178, 1955.
- [47] Y. Lim and Y. Teh. Variational Bayesian approach to movie rating prediction. *Proceedings of KDD cup and workshop*, Vol. 7, pp. 15–21, 2007.
- [48] B. Madore and W. Freedman. Computer Simulations of the Belousov-Zhabotinsky Reaction. *Science*, Vol. 222, No. 4624, 1983.
- [49] S. Maerivoet and B. De Moor. Cellular automata models of road traffic. *Physics Reports*, Vol. 419, No. 1, pp. 1–64, 2005.
- [50] M. Markus and B. Hess. Isotropic cellular automaton for modelling excitable media. *Nature*, Vol. 347, No. 6288, pp. 56–58, sep 1990.
- [51] R. Mu and T. Yamamoto. An Analysis on Mixed Traffic Flow of Conventional Passenger Cars and Microcars Using a Cellular Automata Model. *Procedia - Social and Behavioral Sciences*, Vol. 43, pp. 457–465, 2012.
- [52] T. Musha and H. Higuchi. Traffic Current Fluctuation and the Burgers Equation. *Japanese Journal of Applied Physics*, Vol. 17, No. 5, pp. 811–816, may 1978.

- [53] T. Nagatani. Modified KdV equation for jamming transition in the continuum models of traffic. *Physica A: Statistical Mechanics and its Applications*, Vol. 261, No. 3, pp. 599–607, 1998.
- [54] K. Nagel and M. Schreckenberg. A cellular automaton model for freeway traffic. *Journal de Physique I*, Vol. 2, No. 12, pp. 2221–2229, dec 1992.
- [55] S. Nakajima, I. Sato, M. Sugiyama, K. Watanabe, and H. Kobayashi. Analysis of Variational Bayesian Latent Dirichlet Allocation: Weaker Sparsity Than MAP, 2014.
- [56] S. Nakajima and M. Sugiyama. Theoretical Analysis of Bayesian Matrix Factorization. *Journal of Machine Learning Research*, Vol. 12, No. Sep, pp. 2583–2648, 2011.
- [57] S. Nakajima and S. Watanabe. Variational Bayes Solution of Linear Neural Networks and Its Generalization Performance. *Neural Computation*, Vol. 19, No. 4, pp. 1112–1153, apr 2007.
- [58] F. Nakamura and K. Yamazaki. Two statistical methods for grouping vehicles in traffic flow based on probabilistic cellular automata. In *2014 Joint 7th International Conference on Soft Computing and Intelligent Systems (SCIS) and 15th International Symposium on Advanced Intelligent Systems (ISIS)*, pp. 956–960. IEEE, dec 2014.
- [59] L. Neubert, L. Santen, A. Schadschneider, and M. Schreckenberg. Single-vehicle data of highway traffic: A statistical analysis. *Physical Review E*, Vol. 60, No. 6, pp. 6480–6490, dec 1999.
- [60] G. F. Newell. Nonlinear Effects in the Dynamics of Car Following. *Operations Research*, Vol. 9, No. 2, pp. 209–229, apr 1961.

- [61] K. Nishinari and D. Takahashi. Analytical properties of ultradiscrete Burgers equation and rule-184 cellular automaton. *Journal of Physics A: Mathematical and General*, Vol. 31, No. 24, pp. 5439–5450, jun 1998.
- [62] K. Nishinari and D. Takahashi. A new deterministic CA model for traffic flow with multiple states. *Journal of Physics A: Mathematical and General*, Vol. 32, No. 1, pp. 93–104, jan 1999.
- [63] N. Pelechano and A. Malkawi. Evacuation simulation models: Challenges in modeling high rise building evacuation with cellular automata approaches. *Automation in Construction*, Vol. 17, No. 4, pp. 377–385, 2008.
- [64] Physics of Traffic. <http://traffic.phys.cs.is.nagoya-u.ac.jp/>.
- [65] L. A. Pipes. An Operational Analysis of Traffic Dynamics. *Journal of Applied Physics*, Vol. 24, No. 3, p. 274, 1953.
- [66] I. Prigogine and F. C. Andrews. A Boltzmann-Like Approach for Traffic Flow. *Operations Research*, Vol. 8, No. 6, pp. 789–797, dec 1960.
- [67] A. Rákos and G. M. Schütz. Bethe ansatz and current distribution for the TASEP with particle-dependent hopping rates. jun 2005.
- [68] A. Reuschel. Fahrzeugbewegungen in der Kolonne. *Osterr. Ing. Archiv.*, Vol. 4, No. 1, pp. 193–215, 1950.
- [69] G. M. Schütz. Critical phenomena and universal dynamics in one-dimensional driven diffusive systems with two species of particles. *Journal of Physics A: Mathematical and General*, Vol. 36, No. 36, pp. R339–R379, sep 2003.
- [70] V. Šmídl and A. Quinn. *The Variational Bayes Approach in Signal Processing*. No. April. Springer Science & Business Media, 2004.

- [71] Y. Sugiyama, M. Fukui, M. Kikuchi, K. Hasebe, A. Nakayama, K. Nishinari, S. Tadaki, and S. Yukawa. Traffic jams without bottlenecks-experimental evidence for the physical mechanism of the formation of a jam. *New Journal of Physics*, Vol. 10, No. 3, p. 033001, mar 2008.
- [72] M. Takayasu and H. Takayasu. $1/f$ NOISE IN A TRAFFIC MODEL. *Fractals*, Vol. 01, No. 04, pp. 860–866, dec 1993.
- [73] T.Q. Tang, C.Y. Li, and H.J. Huang. A new car-following model with the consideration of the driver’s forecast effect. 2010.
- [74] M. Treiber, A. Hennecke, and D. Helbing. Congested Traffic States in Empirical Observations and Microscopic Simulations. *Physical Review*, Vol. 62, No. 2, pp. 1805–1824, aug 2000.
- [75] M. Treiber and A. Kesting. *Traffic Flow Dynamics*. Springer, 2013.
- [76] K. Watanabe. An alternative view of variational Bayes and asymptotic approximations of free energy. *Machine Learning*, Vol. 86, No. 2, pp. 273–293, feb 2012.
- [77] K. Watanabe, M. Shiga, and S. Watanabe. Upper bound for variational free energy of Bayesian networks. *Machine Learning*, Vol. 75, No. 2, pp. 199–215, may 2009.
- [78] K. Watanabe and S. Watanabe. Lower bounds of stochastic complexities in variational bayes learning of gaussian mixture models. In *IEEE Conference on Cybernetics and Intelligent Systems, 2004.*, Vol. 1, pp. 99–104. IEEE, 2004.
- [79] K. Watanabe and S. Watanabe. Stochastic Complexities of Gaussian Mixtures in Variational Bayesian Approximation. *Journal of Machine Learning Research*, Vol. 7, No. Apr, pp. 625–644, 2006.

- [80] K. Watanabe and S. Watanabe. Stochastic complexities of general mixture models in variational Bayesian learning. *Neural Networks*, Vol. 20, No. 2, pp. 210–219, 2007.
- [81] S. Watanabe. Algebraic Analysis for Nonidentifiable Learning Machines. *Neural Computation*, Vol. 13, No. 4, pp. 899–933, apr 2001.
- [82] S. Watanabe. *Algebraic geometry and statistical learning theory*. Cambridge University Press, 2009.
- [83] W. G. Weng, T. Chen, H. Y. Yuan, and W. C. Fan. Cellular automaton simulation of pedestrian counter flow with different walk velocities. *Physical Review E*, Vol. 74, No. 3, p. 036102, sep 2006.
- [84] S. Wolfram. Cellular automata as models of complexity. *Nature*, Vol. 311, No. 5985, pp. 419–424, 1984.
- [85] S. Wolfram. Universality and complexity in cellular automata. *Physica D: Nonlinear Phenomena*, Vol. 10, No. 1, pp. 1–35, 1984.
- [86] S. Yukawa and M. Kikuchi. Coupled-Map Modeling of One-Dimensional Traffic Flow. *Journal of the Physical Society of Japan*, Vol. 64, No. 1, pp. 35–38, jan 1995.
- [87] 国土交通省 交通流対策資料. http://www.meti.go.jp/committee/sankoushin/sangyougijutsu/chikyu_kankyo/yakusoku_souan_wg/pdf/005_07_00.pdf.
- [88] 三村昌泰. 現象数理学入門. 東京大学出版, 2013.
- [89] 山崎啓介. 多種粒子 TASEP を表現する混合分布モデルと統計的粒子クラスタリングについて (応用). 日本応用数理学会論文誌, Vol. 24, No. 4, pp. 357–372, 2014.

- [90] 小林浩一, 山崎啓介. 基本図の線形性に基づく SOV モデルのパラメータ推定 (応用). 日本応用数学会論文誌, Vol. 22, No. 4, pp. 287–300, 2012.
- [91] 西山悠, 渡辺澄夫. 完全 2 部グラフ型ボルツマンマシンの平均場近似による確率的複雑さについて. 電子情報通信学会論文誌 A, Vol. J89-A, No. 8, pp. 671–678, 2006.
- [92] 竹内伝史, 本多義明, 青島縮次郎, 磯部友彦. 交通工学. 鹿島出版会, 2000.
- [93] 中村文士, 山崎啓介. 交通流映像からの速度決定則のモデル化とグループ分け (画像処理及び一般). 電子情報通信学会技術研究報告. ITS, Vol. 114, No. 459, pp. 155–160, 2015.
- [94] 中村文士, 山崎啓介. 変分ベイズ法による多種粒子 ZRP の統計的推測について. 日本応用数学会論文誌, Vol. 26, No. 3, pp. 268–285, 2016.
- [95] 中村文士, 渡辺澄夫. 変分ベイズ学習の局所解の構造について (ポスターセッション, 第 14 回情報論的学習理論ワークショップ). 電子情報通信学会技術研究報告. IBISML, 情報論的学習理論と機械学習, Vol. 111, No. 275, pp. 5–10, 2011.
- [96] 中村文士, 渡辺澄夫. 変分ベイズ法の局所解における自由エネルギーと汎化誤差の関係 (ポスターセッション, 第 15 回情報論的学習理論ワークショップ). 電子情報通信学会技術研究報告. IBISML, 情報論的学習理論と機械学習, Vol. 112, No. 279, pp. 59–65, 2012.
- [97] 中村文士, 渡辺澄夫. 一般ディリクレ分布を用いた混合正規分布の変分自由エネルギーについて. 電子情報通信学会技術研究報告. IBISML, 情報論的学習理論と機械学習, Vol. 112, No. 454, pp. 9–16, 2013.
- [98] 中村文士, 渡辺澄夫. 一般ディリクレ分布を用いた混合正規分布の変分自由エネルギーの漸近挙動について. 電子情報通信学会論文誌 D, Vol. J97-D, No. 5, pp. 1001–1013, 2014.

- [99] 中村文士, 渡邊澄夫. ハミルトニアンモンテカルロ法を用いた神経回路網の学習と汎化誤差の推定について. 電子情報通信学会技術研究報告. IBISML, 情報論的学習理論と機械学習, Vol. 116, No. 300, pp. 137–142, 2016.
- [100] 中村文士, 渡邊澄夫. 変分ベイズ法を用いた混合区分線形モデルによる車両の車間-速度特性の推定 (ポスターセッション, 第19回情報論的学習理論ワークショップ). 電子情報通信学会技術研究報告. IBISML, 情報論的学習理論と機械学習, pp. 137-142, 2016.
- [101] 中島伸一. 変分ベイズ学習. 講談社, 2016.
- [102] 中野修弘, 渡邊澄夫. ベイズ事後分布実現における平均場近似の精度評価. 電子情報通信学会技術研究報告. NC, ニューロコンピューティング, Vol. 104, No. 760, pp. 111–116, 2005.