統計力学とデータサイエンス

2025. 01. 29

鈴木　貴（大阪大学）

5. データサイエンスと統計力学の基礎概念

機械学習とデータエンジニアリングは社会の基盤技術となっている

ソリューション企画

要求処理

付加価値連鎖図

引合受付 → 提案準備 → 価格算出 → 要求支援

要求計画 → 要求割当 → 要求管理

組織図

経営管理 ← A

オフィスサービス ← B

C

販売支援 ← D

算出サービス ← E

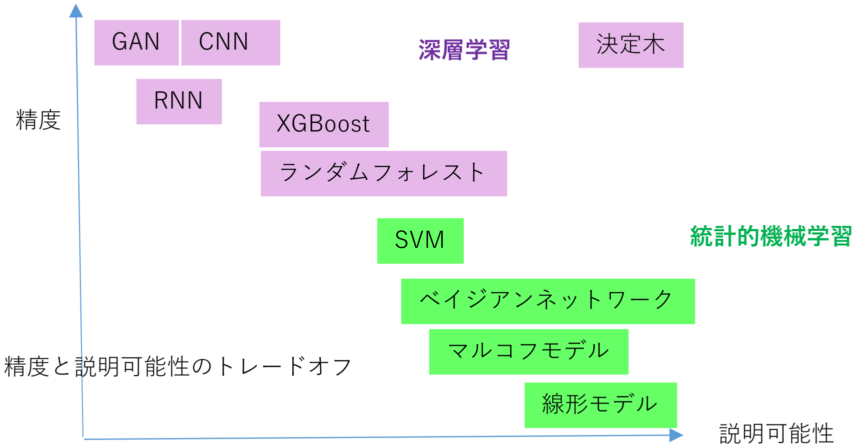
イベント駆動プロセス連鎖図

データエンジニアリング

機械学習

回帰問題
- 重回帰分析
- 情報量基準
- ガウス過程回帰
- 隠れマルコフ
- カルマンフィルタ

分類問題
- 主成分分析
- クラスタリング
- ベイジアンネットワーク
- サポートベクターマシン(SVM)
- スパースモデリング



精度と説明可能性のトレードオフ

データベース

データベース構築はビジネスルール設計そのものである

マルチメディア

- 画像処理
- テキスト処理
- 音声処理

要件定義 → 基本設計 → 詳細設計 → 実装 → 単体テスト → 結合テスト → システムテスト

ウォ-タ-フォ-ルモデル

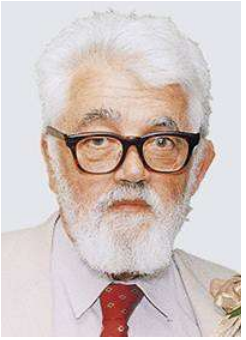
水が流れるように時系列の工程を設計する

スプリント → ソフトウェア → スプリント → ソフトウェア → スプリント → ソフトウェア → スプリント → ソフトウェア →

アジャイル開発

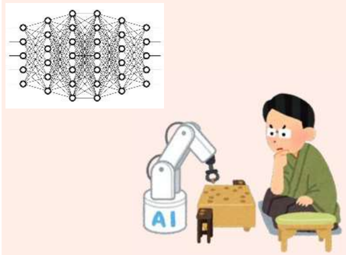
新しいソフトから得られた気付きをもとに要件定義を見直す（フィートバックループ）

機械学習の概要と未来：データに潜む価値



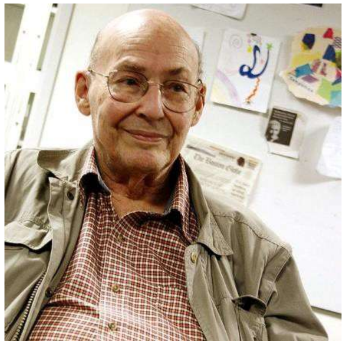
ジョン・マッカーシー
（1927～2011）

人工知能ソフトの プログラミング言語 の開発

制御工学	人工知能		
レベル1	レベル2	レベル3	レベル4
単純な制御プログラム	古典的な人工知能	機械学習	ディープラーニング
✓ あらかじめ決められたルールに基づいて動く	✓ 探索・推論，知識データを利用 ✓ パターンが多いもの ✓ 特定分野で有用	✓ 自動的に判断できる，ルールや知識を自ら学習 ✓ 多くのサンプルデータから学習 ✓ パターン認識から発展	✓ 判断に必要な特徴を学習 ✓ 画像認識，音声認識，自動翻訳
			

ハーバート・サイモン
（1927～2016）

コンピュータ科学と心理学が専門



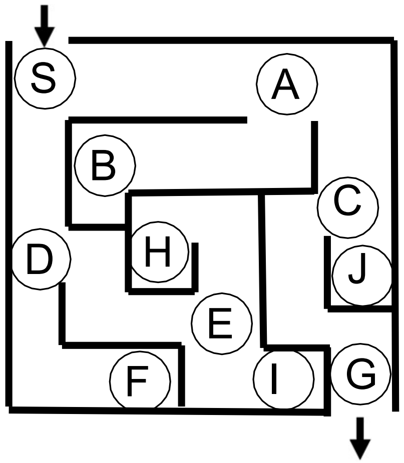
マービン・ミンスキー
（1927～2016）

人工知能最先端の研究所を設立．
「人工知能の父」

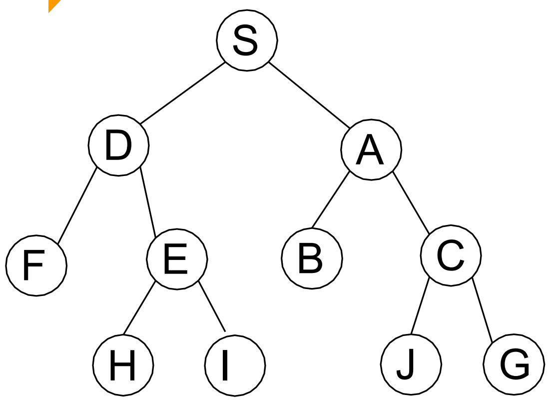
アレン・ニューウェル
（1927～1992）

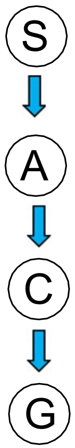
世界初の人工知能プログラム

第1次AIブーム



探索





簡単な問題に限られ，現実（複雑）な問題を解くのは難しく，　限界を迎える

（参考）人間対AI

- チェス，将棋，囲碁で人間が敗れる

	チェス（約10^{120}通り）	将棋（約10^{220}通り）	囲碁（約10^{360}通り）
人工知能	DeepBlue	Ponanza	AlphaGO（碁）
開発者	IBM	山本一成	DeepMind社
対戦年	1997年	2013年	2016年
結果	ゲイリー・カスパロフ（世界チャンピオン）に勝利	現役プロ棋士に勝利	韓国のイ・セドル（世界トップクラスの棋士）に勝利

モンテカルロ法

- ランダムな手を打ち，ゲームを終局させる（プレイアウト）
- プレイアウトを複数回実行して，勝率に応じてゲームの スコアを評価
- 人間がスコアを考えるより，優れている
- ブルートフォース（力任せ）
  - 全ての組み合わせを隈なく調べる
- AlphaGo：　ディープラーニングの技術でプロ棋士に勝利

定理　全てのボードゲームは先手必勝、後手必勝、千日手のいずれかに分類される

第2次AIブーム：コンピュータに知識をもたせる

- 現実の産業にコンピュータを役立てる試み
- 専門家（エキスパート）の知識を取り込み活かす
- エキスパートシステム
  - MYCN(マイシン）：病気を診断するエキスパートシステム
- エキスパートシステムの課題
  - 専門家からたくさんの知識を取り出す
  - 膨大な知識を整理
  - あいまいなデータを理解できない
- 知識を表現する
  - オントロジー：概念の意味や関係を文字や記号で表す学問

知識ベースとエキスパートシステム

- マイシン（MYCIN）
  - 1970年代にスタンフォード大学で開発
  - 血液中のバクテリアの診断支援
  - ルールベース
  - 質問に順に回答すると，感染した細菌を特定
  - 精度：専門医 ＞ マイシン（MYCIN）＞ 専門外の医師
- DENDRAL
  - 1960年代にエドワード・ファイゲンバウムがスタンフォー ド大学で開発
  - 1977年：実世界の問題に対する技術を重視する「知識工学」を提唱

人工無能

- チャットボット，おしゃべりボット
  - コンピュータプログラム
  - 特定のルール・手順に沿って会話を機械的に処理
  - 会話内容は理解していない
- イライザ（ELIZA）
  - 人工無能の元祖
  - 1964年から1966年にジョセフ・ワイゼンバウムにより開発
  - あらかじめ用意されたパターンと比較
  - パターンに応じた返答
  - 相手の発言を再利用（オウム返し）

知識獲得のボトルネック

- 専門家，ドキュメント，事例などから知識を獲得
- ドキュメントや事例からは自然言語処理や機械学習の利用
- 専門家の知識：
  - 経験的
  - 暗黙的
  - 獲得が困難
- 知識ベースの保守が困難
  - 大量の知識の一貫性などをチェック
- 常識を扱うには？

データと機械学習（次のブームのきっかけ）

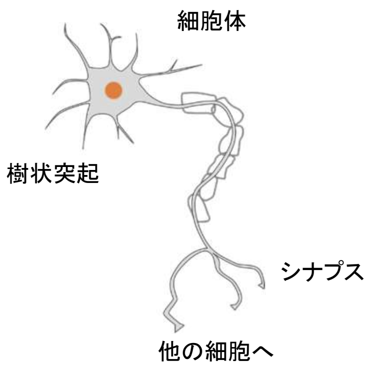
- 機械学習
  - コンピュータがサンプルデータにより学習
  - サンプルデータが多いほど良い学習結果が得られるビッグデータ
  - インターネットの普及
  - 1990年Webの登場で，大量のデータを入手可能

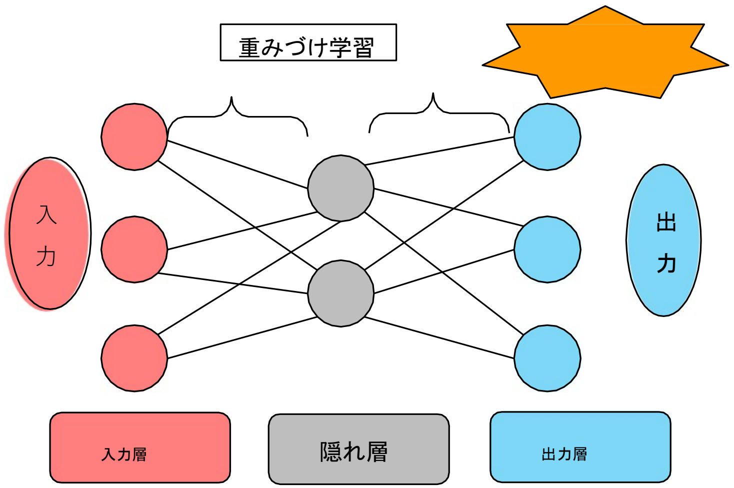
機械学習と統計的自然言語処理

- 従来：文法構造や意味構造を分析して単語単位で訳を割り当てる
- 統計的自然言語処理
- 複数の単語をひとまとまりにした単位で用意された膨大な量の対訳データ
- 最も正解である確率の高い訳を選択
- コーパス
- 自然言語の文章を構造化し，品詞などの情報を付与して大規模に集積したもの
- コンピュータ利用のため，電子化データで提供されている

第3次人工知能ブーム：ディープラーニングの登場

- 人間の脳をまねて「分ける」：ニューラルネットワーク
- 「ニューロン」と呼ばれる神経細胞
- たくさんのニューロンで電気刺激（電気信号）をやり取り





ニューラルネットワーク

ディープラーニングによる技術の発展

- 学習データが十分でないものや説明が必要なものには 弱い
- 現状では最強のツール
- 画像，音声，自然言語処理

データサイエンスとは何か：相関と因果を中心として

1つのベクトル
統計量を持つ

1変量データ

平均 ＝ 71

ID	Point	**Square variance from the mean**
1	70	1
2	65	36
3	100	841
4	80	81
5	75	16
6	90	361
7	95	576
8	30	1681
9	50	441
10	55	256

分散 $s^2 = \dfrac{1+36+841+81+16+361+576+1681+441+256}{10} = 429$

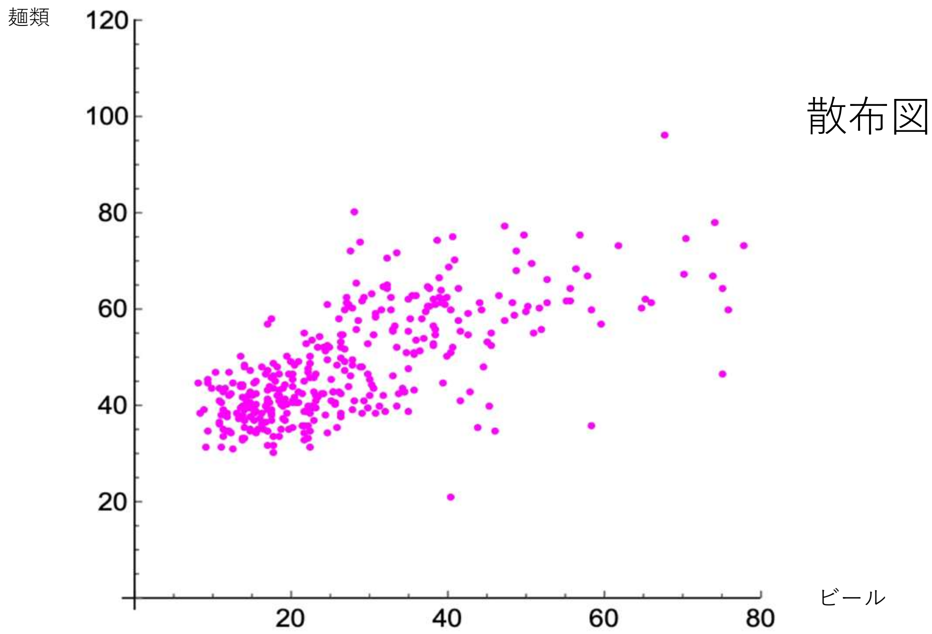
標準偏差 $s = \sqrt{429} = 20.71232$

2変量データ

同一次元の二つのベクトル
個別の他に全体としての統計量を持つ
様々な表示（可視化）ができる

ビール　麺類

$(x_1, y_1), \cdots, (x_N, y_N),\ N = 365$ days



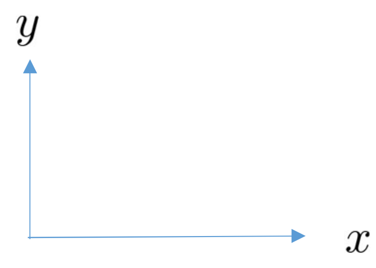
相関係数 $r = \dfrac{s_{xy}}{s_x s_y}$

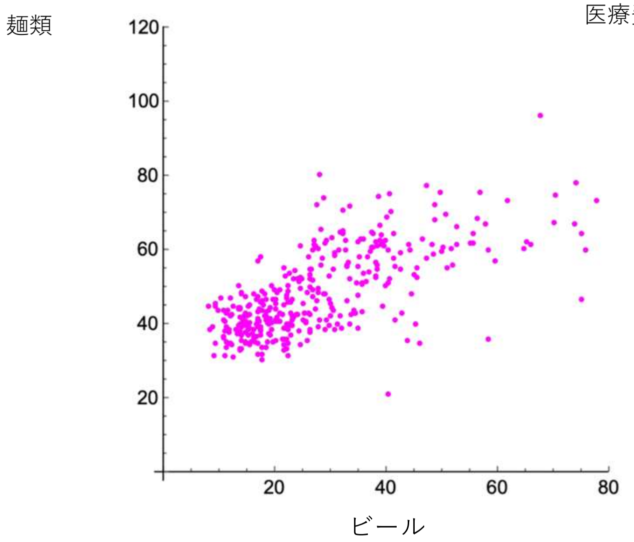
相関と因果

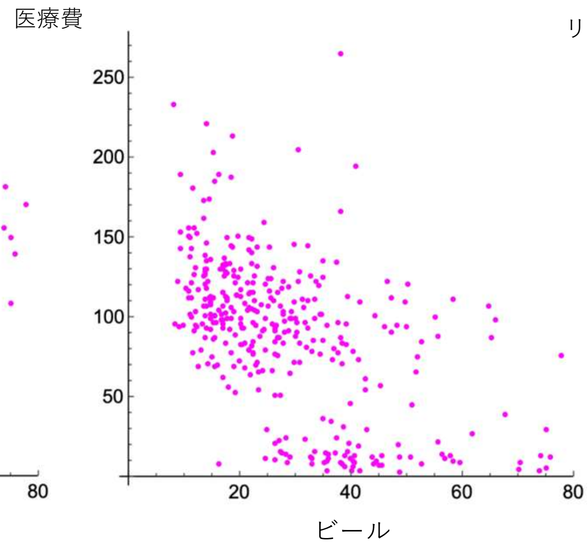
ビール　麺類

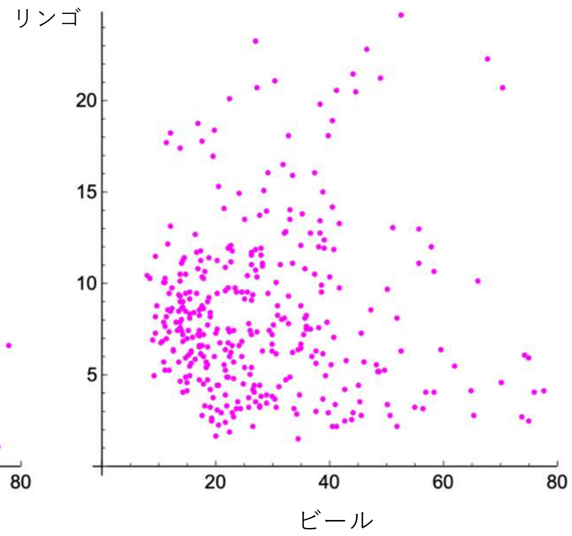
$$(x_1, y_1), \cdots, (x_N, y_N),\ N = 365\ \text{days}$$

毎日の支出（1万世帯合計）の表示









正の相関がある

負の相関はあっても因果ではない

交絡因子

相関がない

次元削減

ゲノム数　細胞数

$$x_i \in \mathbf{R}^N,\ 1 \le i \le m$$

ビッグデータをどう視覚化するか

$$w \in \mathbf{R}^N,\ |w| = 1$$

$$y_i = w^T x_i,\ 1 \le i \le m$$

1次元空間への射影

(N, N) 行列

共分散行列

主成分分析

$$\sigma^2 = \frac{1}{m}\sum_{i=1}^{m}(y_i - \overline{y})^2 = \frac{1}{m}\sum_{i=1}^{m}(w^T x_i - w^T \overline{x})(x_i^T w - \overline{x}^T w) = w^T \left[\frac{1}{m}\sum_{i=1}^{m}(x_i - \overline{x})(x_i - \overline{x})^T\right] w = w^T S w$$

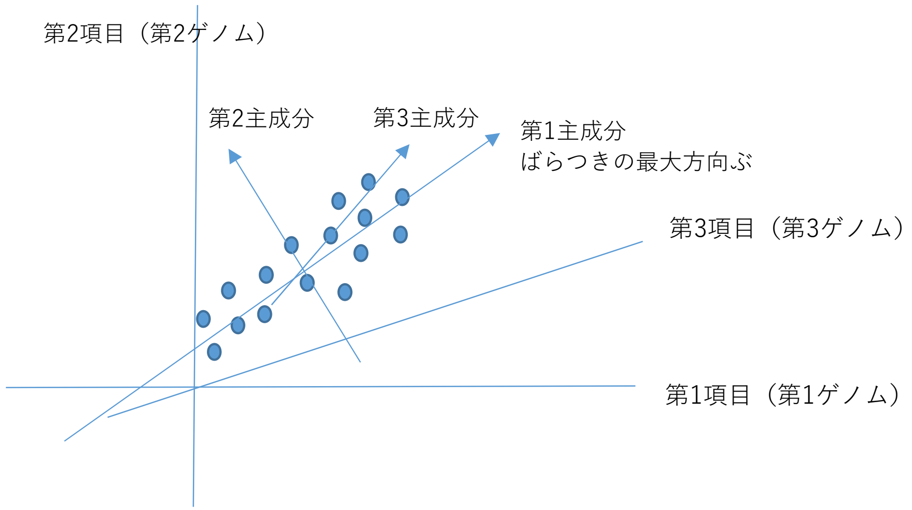
最大化

第1主成分の検出

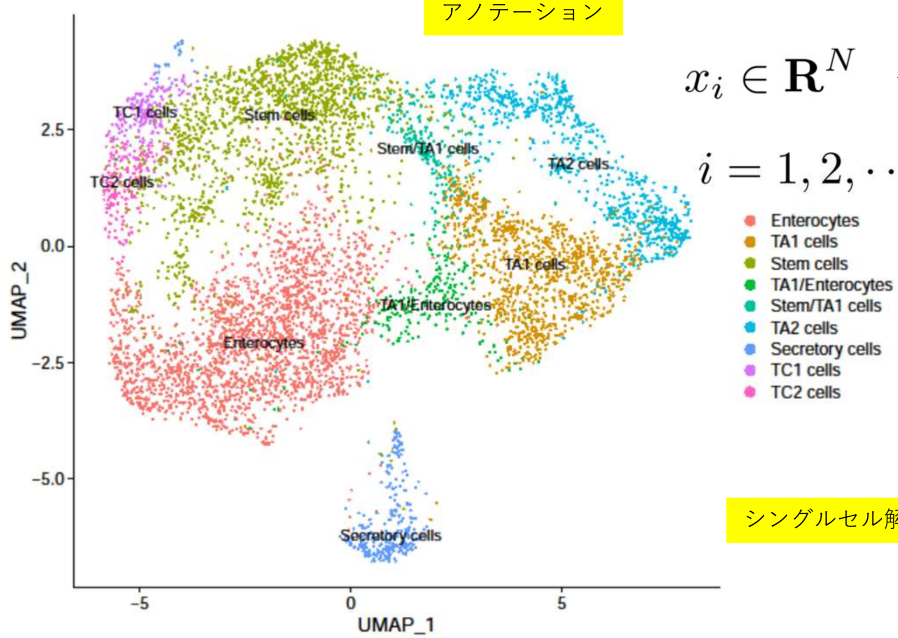
2次形式・対称行列・対角化

$$\overline{y} = \frac{1}{m}\sum_{i=1}^{m} y_i = w^T \overline{x} \in \mathbf{R}$$

$$\overline{x} = \frac{1}{m}\sum_{i=1}^{m} x_i \in \mathbf{R}^N$$



アノテーション



$$x_i \in \mathbf{R}^N$$ ゲノム発現

$$i = 1, 2, \cdots$$ 細胞

シングルセル解析

特異ベクトルを用いたクラス分け

品目

A

都道府県

$m > n$ 縦長

家計調査の				SSDSE-2020C の　情　報		備考
						の
外食			1.12	LB12	12 外 食	
	外　食	食　代	390	LB121101		
			391	LB121102	中　華	
			392	LB121103	他の麺類外食	
			393	LB121104	すし（外食）	
			394	LB121105	和食	
			39A	LB121106	中華食	
			395	LB121107	洋食	
			399	LB121108	焼肉	
			39B	LB121109		
			396	LB121110	他の主食的外食	
			397	LB121201	喫茶代	
			398	LB121202	飲酒代	
	学校給食		39X	LB122001	学校給食	

列：10品目
1. 日本そば・うどん 2. 中華そば 3. すし（外食）4. 和食 5. 中華食
6. 洋食 7. 焼肉 8. ハンバーガー 9. 喫茶代 10. 飲酒代

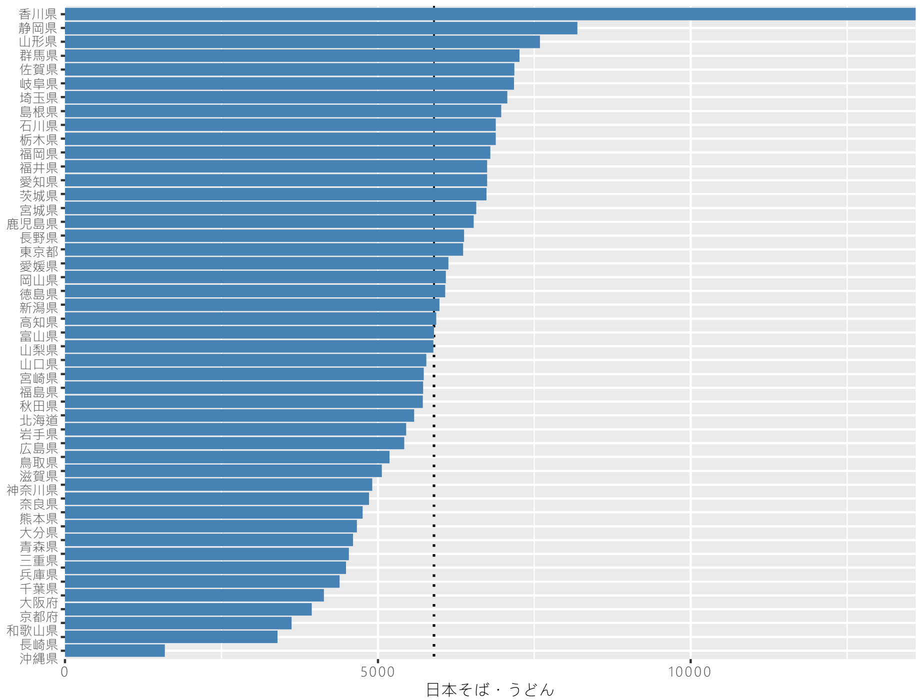
行：47都道府県

都道府県庁所在市別・二人以上の世帯1 世帯当たり・ 品目別の年間支出金額（2019年～2021年の平均）

2次データ

例：日本そば・うどん

列ベクトルを成分の大きさでソート



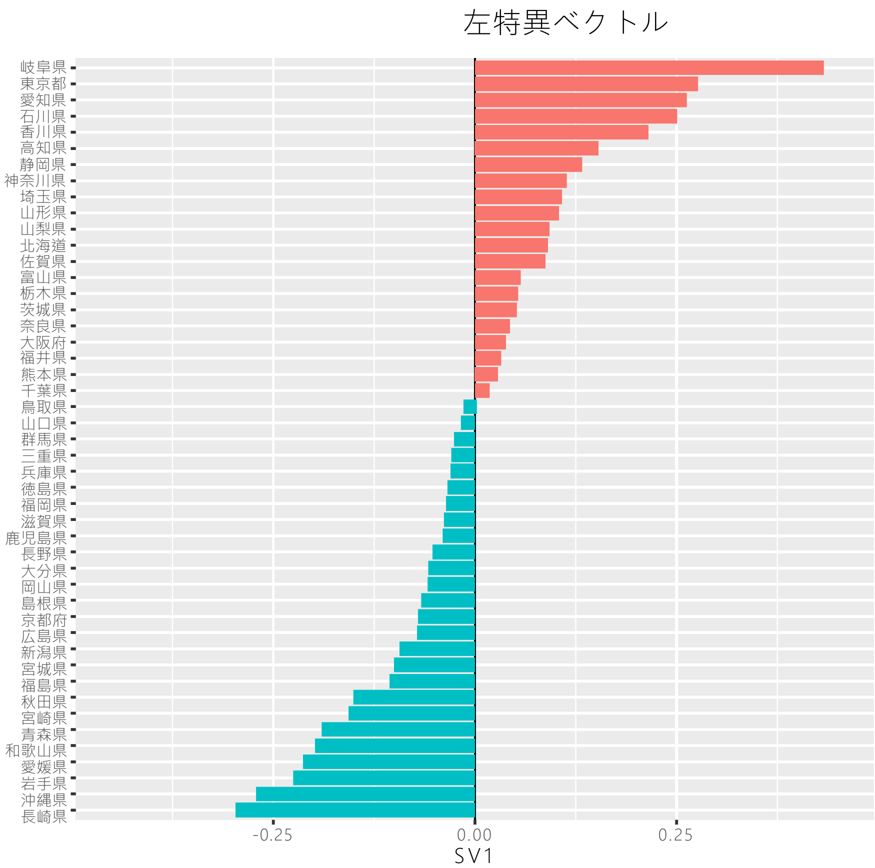
都道府県ベクトル　品目ベクトル

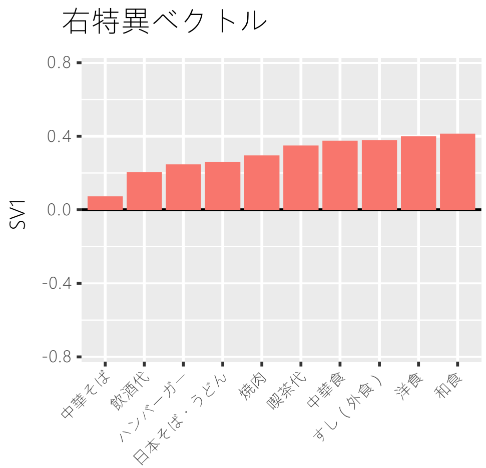
$$A = \sigma_1 u_1 v_1^T + \cdots + \sigma_r u_r v_r^T$$

$$\sigma_1 \geq \cdots \geq \sigma_r > 0$$

第1 潜 在 変 数

成分の大きさでソート





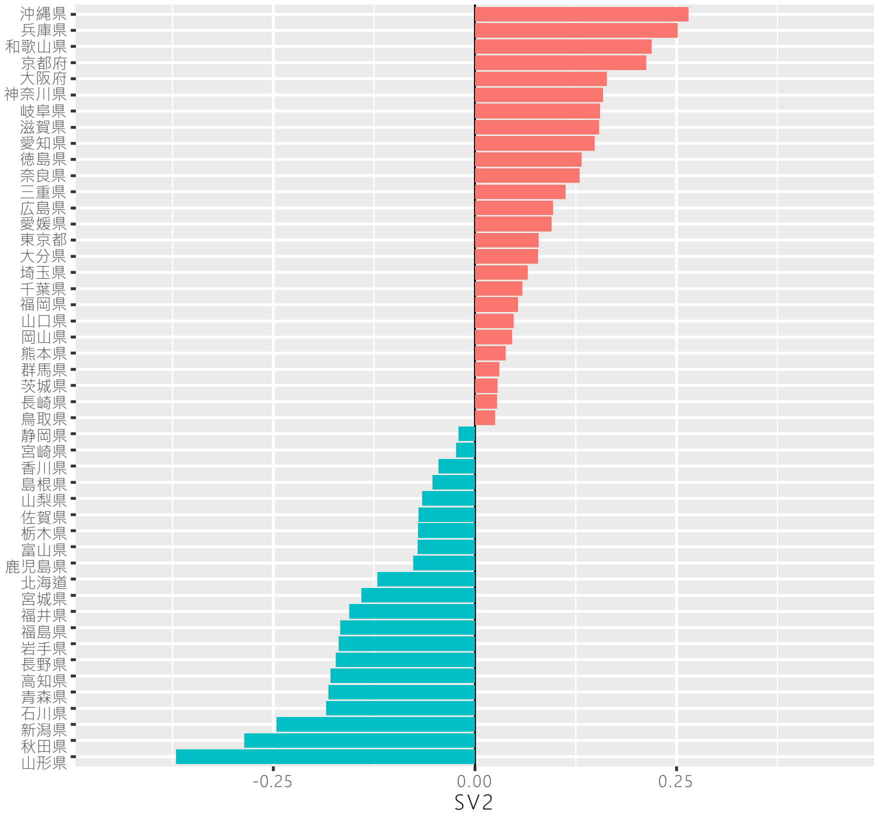
外食費全般

品目をアノテーション

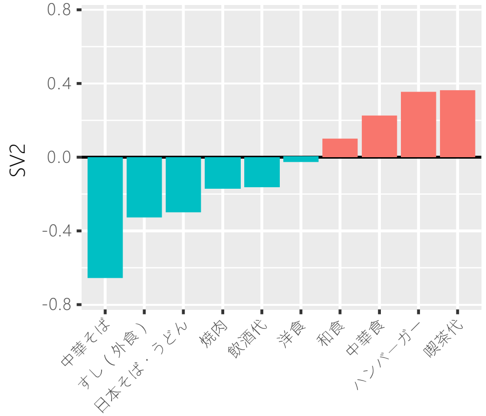
第2潜在変数

成分の大きさでソート

左特異ベクトル



右特異ベクトル

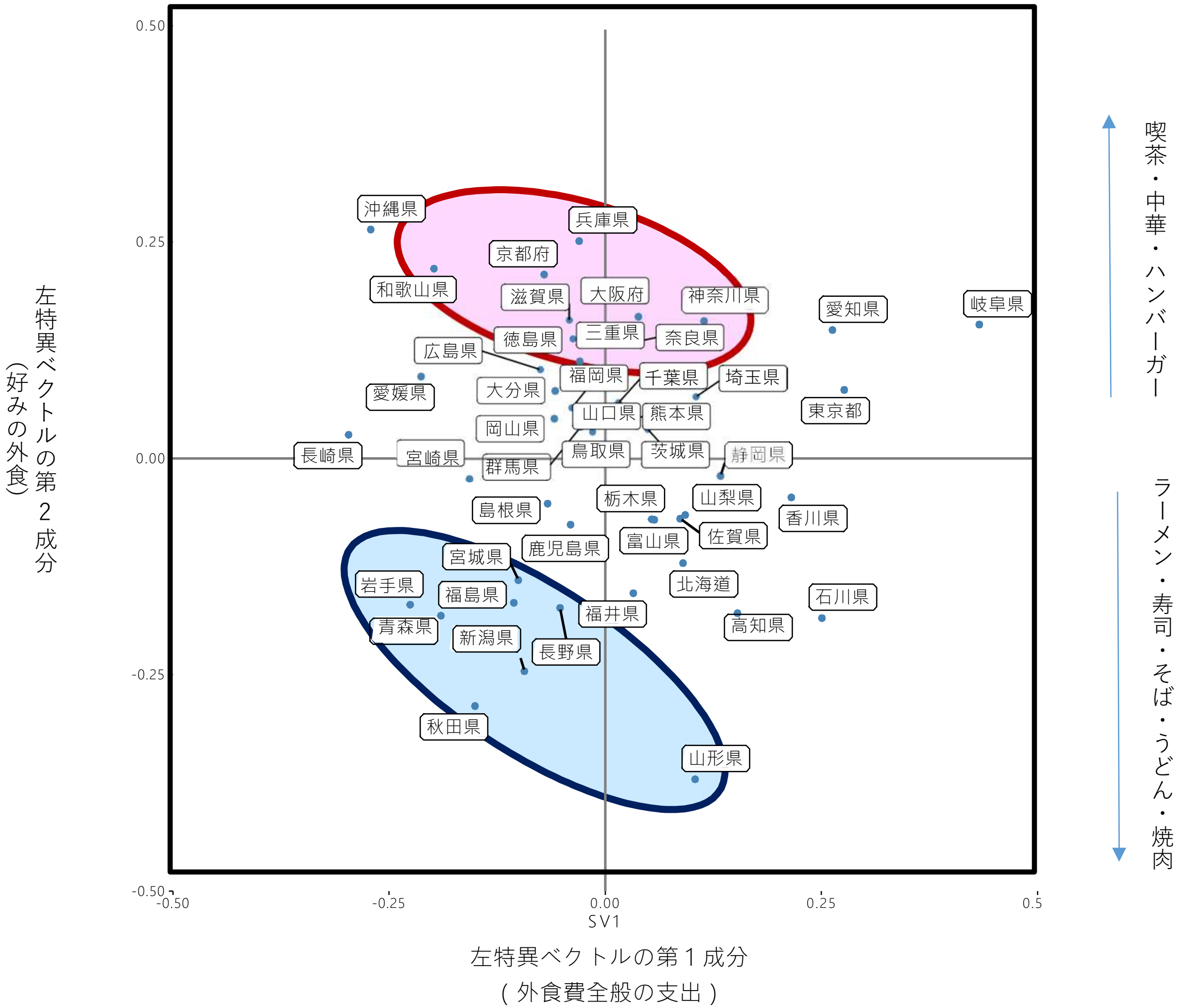


麺類×ハンバーガー/喫茶

品目をアノテーション

食費支出の層別化

行のパターン（都道府県）を第1・第2品目で散布表示 → 外食費のクラス分け（層別化）



どこまで信用してよいか
意思決定に関わるのは問題ないか

ビッグデータとスモールデータ：記述統計と推測統計の世界

エクセルを用いた回帰分析

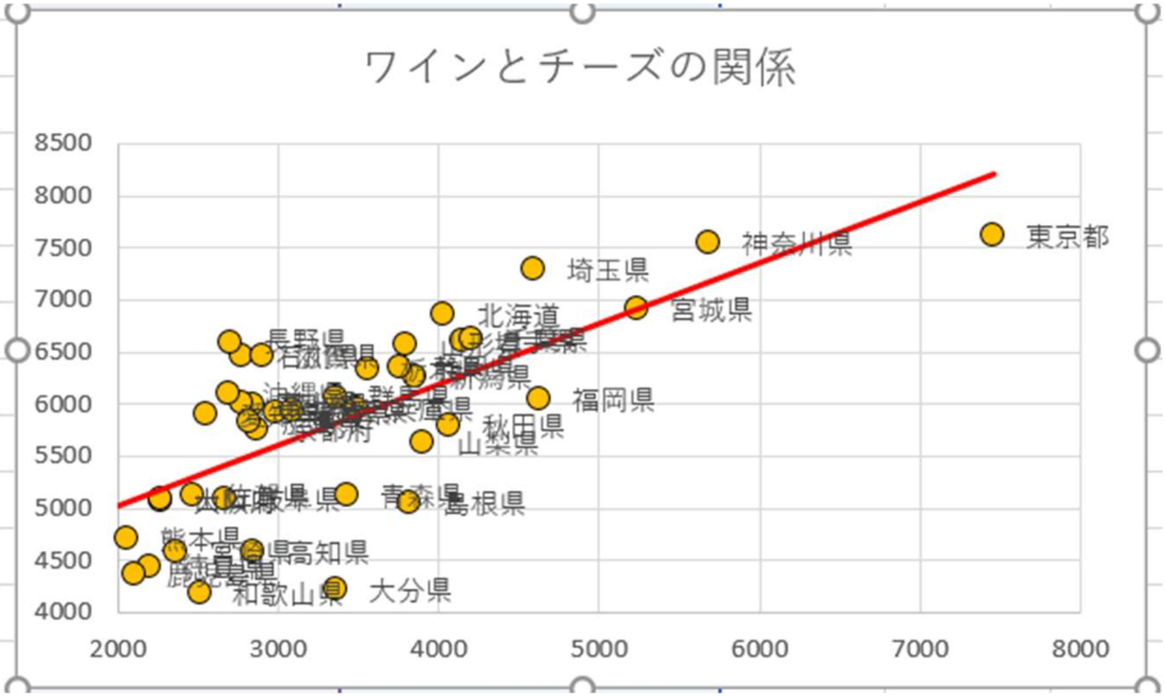
因果関係の推定が適切であるかどうかを**検定**する

p値を用いて説明関数の統計的有意性を検証する

帰無仮説「説明変数が目的変数に影響を与えない」 $\xrightarrow{\text{統計的仮説検定}}$ $p < \underset{\text{有意水準}}{0.05}$ その説明変数が目的変数に影響を与えていないとは言えない

モデルの正当性を**F値**によって検証する $\xrightarrow{\text{分散分析}}$ $F < \underset{\text{有意水準}}{0.05}$ モデルは母集団の説明になっていないとは言えない

	A	B	C	D	E	F	G	H	I
1	概要								
2									
3	回帰統計								
4	重相関 R	0.749259							
5	重決定 R2	0.561389							
6	補正 R2	0.551642							
7	標準誤差	605.6678							
8	観測数	47							
9									
10	分散分析表								
11		自由度	変動	分散	された分散	有意 F			
12	回帰	1	21128335	21128335	57.59652	1.37E-09			
13	残差	45	16507508	366833.5					
14	合計	46	37635843						
15									
16		係数	標準誤差	t	P-値	下限 95%	上限 95%	下限 95.0%	上限 95.0%
17	切片	3852.229	256.8694	14.99684	4.03E-19	3334.867	4369.59	3334.867	4369.59
18	X 値 1	0.58301	0.076821	7.589237	1.37E-09	0.428286	0.737735	0.428286	0.737735
19									



データサイエンスは統計学と情報科学が融合する実学である

統計学 → データサイエンス

情報科学

ゴルトン
ピアソン
フィッシャー

進化・遺伝・DNA

2nd generation

1

2

回帰

1st generation

・線形代数
・集合と位相
・実解析
・確率論
・最適化
・グラフ理論
・統計力学
・微分幾何
・代数幾何
・代数的位相幾何

第1次AIブーム

第2次AIブーム

第3次AIブーム

1960

チューリングテスト

ダートマス会議（1956）
デジタル計算機
パーセプトロン
チューリングテスト

XORゲート

1980

パーセプトロン

エキスパートシステム
知識工学
多層化
バックプロパゲーション

勾配消失

計算資源
アルゴリズム

2000

深層学習

ビッグデータ
深層学習
オートエンコーダ
畳み込みネットワーク

ブラックボックス

ビッグデータ
機械学習

2020

生成AI

予測
分類・識別
データ生成

Ethical-Legal-Social Issue

公開データベース
共有/フィードバック

2040

シンギュラリティ

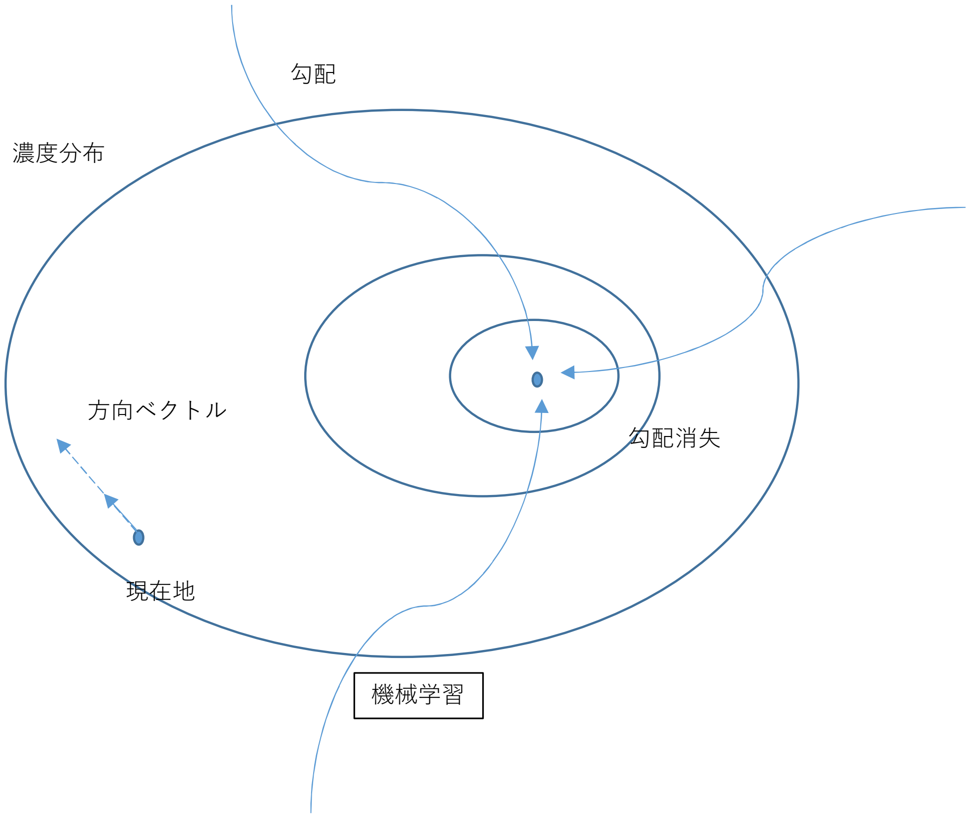
説明可能AI

電力問題

量子コンピュータ
核融合発電
スピントロニクス

統計学は記述統計と推測統計に分かれる

推測統計は母集団の統計量を推定・検定する手法であり頻度論統計とベイズ統計に分かれる



データサイエンスで使われる基本的な数学は４つ

線形代数　**微積分**　**最適化**　**確率論**

モデル
スキーム
アルゴリズム

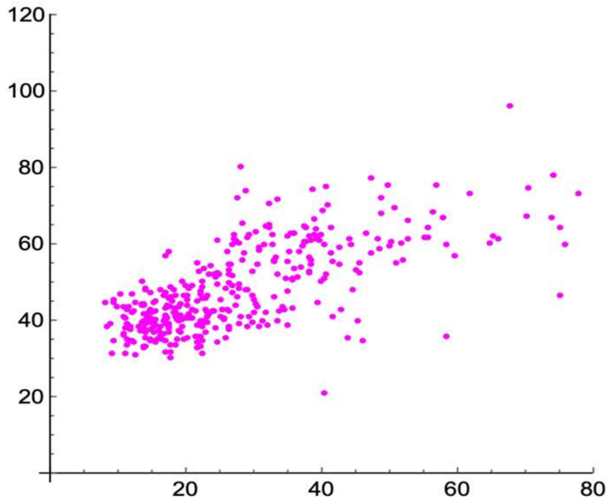
機械学習は深層学習と統計的機械学習に分かれる

推測統計と統計的機械学習は現代数学を特に活用する

連立1次方程式

$$
\begin{aligned}
a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + \cdots + a_{2n}x_n &= b_2 \\
\vdots \quad & \quad \vdots \\
a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m
\end{aligned}
$$

ビッグデータ・スモールデータ分析



$\boxed{\text{分布}}$ 1変量データを見直す $y = (y_j) \in \mathbf{R}^N$ 1変量データ $\mu = \frac{1}{N}\sum_{j=1}^{N} y_j$ 平均 $\sigma^2 = \frac{1}{N}\sum_{j=1}^{N}(y_j - \mu)^2$ 分散

データの値の集合（階級）

$Z \equiv \{z_k \mid 1 \le k \le n\} = \{y_j \mid 1 \le j \le N\},\ z_k \ne z_{k'},\ k \ne k'$ 分布

個数

$J_k = \sharp\{j \mid y_j = z_k\}$ $z_k \in Z$ の度数

記述統計
ベクトル（データ）

$\Longrightarrow$

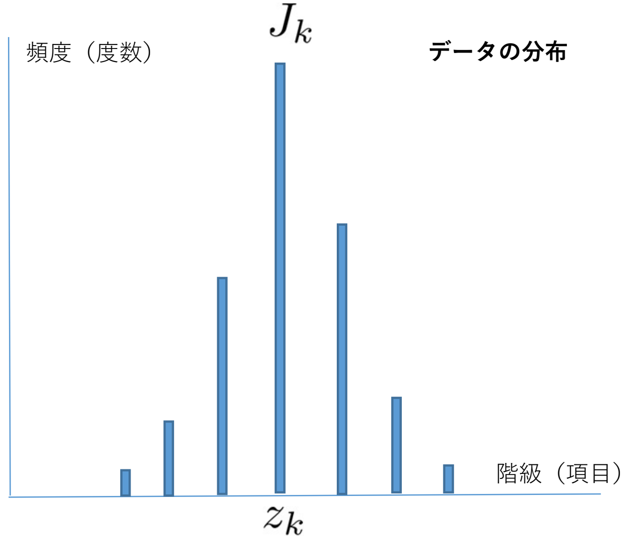
推測統計
確率変数（標本）

確率 事象

母集団は**確率空間** $\Omega = \{1, 2, \cdots, N\}$ $P(\{j\}) = \frac{1}{N},\ 1 \le j \le N$

母集団から無作為に標本を1つ選ぶ操作は**確率変数** $X : j \in \Omega \mapsto y_j \in Z$

$$E[X] = \sum_j y_j P(\{j\}) = \mu, \quad V[X] = \sum_j (y_j - \mu)^2 P(\{j\}) = \sigma^2$$



1変量データの統計量（平均・分散）はその分布（階級と度数）によって定まる
1変量データ（ベクトル）の平均・分散は標本（確率変数）の期待値・分散と一致する

多変量データの標本は同一確率空間上のベクトル値の確率変数 身長・体重・・・

標本抽出

母数

標本抽出の目的は母集団の統計量を推定・検定することである

記述統計には推定と仮説検定がある

推定には点推定と区間推定がある

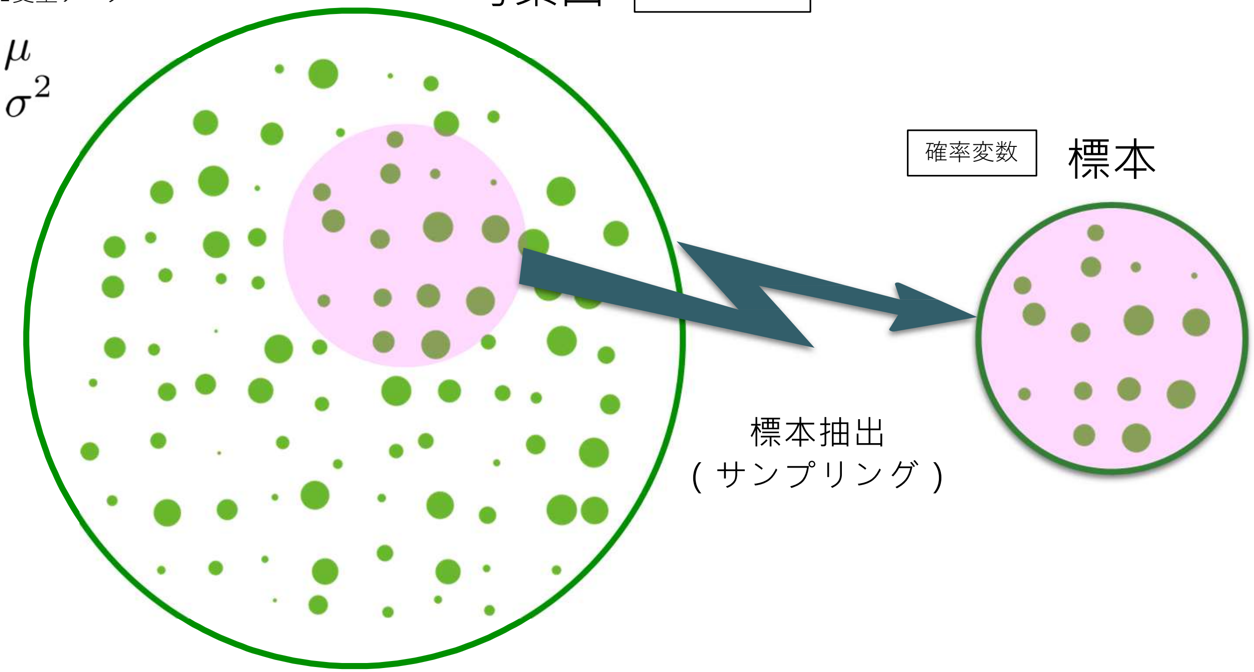
$x_1, x_2, \cdots, x_n$ 標本値

標本は母集団の分布に従う確率変数である
無作為抽出は独立同分布の確率変数を生成する

統計量　1変量データ

平均 μ
分散 σ^2

母集団　標準確率空間



$X_1, X_2, \cdots, X_n$ **標本** **独立同分布**

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$ 標本平均

期待値

$$E[\overline{X}_n] = \sum_{i=1}^{n} E\left[\frac{X_i}{n}\right] = n \cdot \frac{\mu}{n} = \mu$$

分散

$$V[\overline{X}_n] = \sum_{i=1}^{n} V\left[\frac{X_i}{n}\right] = n \cdot \frac{\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

チェビシェフの不等式　確率収束

$$\forall \varepsilon > 0,\ E[|\overline{X}_n - \mu| > \varepsilon] \leq \frac{1}{\varepsilon^2} V[\overline{X}_n] = \frac{\sigma^2}{n\varepsilon^2} \to 0,\ n \to \infty$$ 大数の法則

サンプル数を多くすると標本抽出のばらつきが緩和されて標本平均の期待値は母平均に近づく

標本平均は母平均を近似する（**一致性**）

点推定の基準

不偏性

確率空間 1変量データ 独立同分布確率変数

$X:\Omega\to\mathbf{R}$ 母集団 $X_i:\Omega\to\mathbf{R},\ i=1,2,\cdots,n$ 標本

平均・分散 点推定 確率変数 推定値は平均的に正確である

θ 母集団の統計量 $\hat{\theta}_n=\hat{\theta}_n(X_1,\cdots,X_n)$ θ の推定値 $\Longrightarrow$ $E[\hat{\theta}_n(X_1,\cdots,X_n)]=\theta$

例 **標本平均** $\overline{X}=\dfrac{1}{n}\sum_{i=1}^{n}X_i$ は母平均 μ の不偏推定量

例 **不偏分散** $S^2=\dfrac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$ は母分散 σ^2 の不偏推量

一致性

「大数の法則」 「チェビシェフの不等式」

$\forall\varepsilon>0,\ \lim_{n\to\infty}P(|\hat{\theta}_n-\theta|>\varepsilon)=0$

標本平均は母平均の一致推定量

不偏分散（標本分散）は母分散の一致推定量

標本数を増やせば推定値は母数に漸近する

一様最小分散性

最小2乗近似

$E[(\hat{\theta}_n-\theta)^2]$ を最小にする $\hat{\theta}_n$ 不偏性のもとで 分散 $V[\hat{\theta}_n]$ を最小にするもの

最もばらつきが少ない推定法を選択する

例 X_1,X_2,X_3 独立同分布 $\overline{X}=\dfrac{1}{3}(X_1+X_2+X_3),\ \tilde{X}=\dfrac{1}{2}(X_1+X_3)$ はともに母平均の不偏推定量

標本平均は母平均の一様最小分散不偏推定量 $V[\overline{X}]=\dfrac{\sigma^2}{3},\ V[\tilde{X}]=\dfrac{\sigma^2}{2}$ σ^2 母分散

独立同分布

$X_i : \Omega \to \mathbf{R},\ 1 \le i \le n$ 標本

同時密度関数　独立性

$$f(x_1, \cdots, x_n) = \prod_{i=1}^{n} f(x_i)$$

母集団（確率空間） $(\Omega, \mathcal{F}, P)$

分布関数　確率変数　密度関数

$$F(a) \equiv P(X \le a) = \int_{-\infty}^{a} f(x)dx$$

重積分

$$P((X_1, \cdots, X_n) \in A) = \int_A f(x_1, \cdots, x_n)\ dx_1 \cdots dx_n$$

$A \subset \mathbf{R}^n$ ボレル集合

最尤推定の思想

一様最小分散推定は存在しなかったり不合理となる例も多い

モデルを定めて観測の独立性を仮定し、得られたデータが最も起こりやすいパラメータを選ぶ

母集団の統計量 θ（母数） を観測値 $x_i,\ 1 \le i \le n$（実現値） から推定するとき $L(\theta) = f(x_1, \cdots, x_n; \theta)$ をその尤度という

点推定

最尤推定

尤度 $L(\theta) = L(x_1, \cdots, x_n; \theta)$ を最大化する $\theta = \theta(x_1, \cdots, x_N)$ を求めて $\hat{\theta}_n = \theta(X_1, \cdots, X_n)$ とする

密度関数の形を決める　正規, ポアソン, …

フィッシャー情報量　クラメール・ラオの不等式

推測統計は頻度論統計とベイズ統計に分かれる

母数

パラメータは定数である（**頻度論統計学**） ↔ 母数は確率変数である（**ベイズ統計学**）

ベイズ推定の思想

同時確率　条件付確率

$$P(A, B) = P(B \mid A)P(A) = P(A \mid B)P(B) \Longrightarrow P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}$$

ベイズの定理

ベイズ更新

事後確率　事前確率

同時分布　条件付分布　母数分布

$$f(x, \theta) = f(x \mid \theta)f(\theta) = f(\theta \mid x)f(x) \Longrightarrow f(\theta \mid x) = \frac{f(x \mid \theta)f(\theta)}{\int f(x \mid \theta)f(\theta)d\theta}$$

データ　母数

統計学では密度を分布と呼ぶことが多い

事後分布　最尤　尤度　事前分布　主観分布

ベイズ　周辺尤度　計算量

ベイズ統計はベイズ更新という利点と事前分布と周辺尤度という難点を持っている

$$f(x) = \int f(x \mid \theta)f(\theta)d\theta$$

EAP（例） $E(\theta \mid x) = \int \theta f(\theta \mid x)d\theta = \int \theta \frac{f(x \mid \theta)f(\theta)}{f(x)}d\theta$ を最小にする θ を選ぶ

KL（カルバック・ライブラー）情報量

$$\int_{\mathbf{R}} g dx = \int_{\mathbf{R}} f dx = 1$$

標準確率空間　確率変数（未知）　未知

$(\Omega, \mathcal{F}, P)$　$X : \Omega \to \mathbf{R}$　$g : \Omega \to \mathbf{R}_+ = [0, \infty)$ 密度関数　$f : \Omega \to \mathbf{R}_+$ そのモデル

$g(x) = g_i,\ i-1 < x \leq i$ ⟹ X 離散的　$P(X=i) = g_i,\ i = 1, 2, \cdots,$

各 i に対して母集団の部分集合 $\{X = i\}$（未知）の各元に f の値 f_i を割り当てる

その時の場合の数 W（weight factor）が最大となる f を選択する（熱平衡）（小正準統計）T.S. Mean Field Theories and Dual Variation, 2nd ed. Atlantis Press, Paris 2015 第9章

分布とみなして

$$W = \prod \frac{f_i^{g_i}}{g_i!} \Rightarrow \underset{\text{ボルツマンエントロピー}}{\log W} = \sum (g_i \log f_i - \log g_i!) \underset{\text{スターリングの公式}}{\sim} \sum g_i (\log f_i - \log g_i) = \sum g_i \log \frac{f_i}{g_i}$$

$$\sim \int_{\mathbf{R}} \left\{ \log \frac{f(x)}{g(x)} \right\} g(x)\ dx = -I(g, f)$$

KL情報量　エンセンの不等式

$$I(g, f) = \int_{\mathbf{R}} \{-\log \frac{f(x)}{g(x)}\} g(x)\ dx \geq -\log \int_{\mathbf{R}} \frac{f(x)}{g(x)} g(x)\ dx = 0$$

等号 ⟺ $f = g,\ \text{a.e.}\ \mu(dx) = g(x) dx$

KL情報量

$$I(g,f) = \int_{\mathbf{R}} \left\{ \log \frac{g(x)}{f(x)} \right\} g(x)\ dx$$

例

真の分布

$$g(x) \sim N(0,1)$$

モデル１　モデル２

$$f_1(x) \sim N(0.5,1),\ f_2(x) \sim N(0,1.5)$$

better

$$\Longrightarrow\ I(f_1,g) = 0.125,\ I(f_2,g) = 0.036$$

一定　平均対数尤度　$E\log f(x)$ 最大化

$$I(g,f) = \int_{\mathbf{R}} \{\log g(x)\} g(x) dx - \int_{\mathbf{R}} \{\log f(x)\} g(x) dx$$

最尤推定の原理

$f = f(x)$　$x_1, \cdots, x_n$ 独立な観測データ $\Longrightarrow$ $E\log f(x) \approx \frac{1}{n}\sum_{i=1}^{n} \log f(x_i)$

$\ell = \sum_{i=1}^{n} \log f(x_i)$ 対数尤度

例

$$f_1(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \Longrightarrow \ell_1 = -5\log 2\pi - \frac{1}{2}\sum_{i=1}^{n} x_i^2$$

正規分布

$$f_2(x) = \frac{1}{\pi(x^2+1)} \Longrightarrow \ell_2 = -10\log \pi - \sum_{i=1}^{n} \log(x_i^2+1)$$

コーシー分布

観測データをみて対数尤度が大きいモデルを選択する

統計量だけでなく分布もモデル化できる

情報量統計では観測データに即して推定することができ標本分布論も統計数値表も不要である

AIC（赤池情報量）

平均対数尤度 $E \log f(x)$ 最大化

$$I(g,f)=\int_{\mathbf{R}}\left\{\log\frac{g(x)}{f(x)}\right\}g(x)\ dx=\int_{\mathbf{R}}\{\log g(x)\}g(x)dx-\int_{\mathbf{R}}\{\log f(x)\}g(x)dx$$

$f=f(x\mid\theta)$ $x_1,\cdots,x_n$ 独立な観測データ $\longrightarrow$ $E\log f(x\mid\theta)\approx\frac{1}{n}\sum_{i=1}^{n}\log f(x_i\mid\theta)=\ell(\theta)$

過学習を排除

定理

想定したモデルの中に
真のモデルが含まれる $\longrightarrow$

対数尤度

$$\max_\theta \ell(\theta)=\max_\theta E\log f(x\mid\theta)+k+O\left(\frac{1}{n}\right)$$

$k=\dim\theta$ パラメータ数

AIC $-2\ell(\theta)+2k$ を最小にする θ を選ぶ

エントロピーや情報量を用いるとモデル選択とパラメータ推定が一挙にできる

相対エントロピーの理論

可測関数（確率変数）

$$f(x)\geq 0,\ g(x)\geq 0,\ f^2(x)+g^2(x)>0 \qquad \Phi(s)=s(\log s-1)+1\geq 0,\ s>0$$

$$E(g\mid f)\stackrel{\text{def}}{=}\int[\Phi\left(g/f\right)(x)]f(x)dx\sim\int g\{\log\frac{g}{f}\}\mu(dx)=I(g,f)$$

relative entropy

Czisza'r-Kullback

$$\longrightarrow \|f-g\|_1\leq CE(g\mid f)$$

logarithmic Soblev

$$E(f|\overline{f})\leq C\int|\nabla\sqrt{f}|^2dx,\ \overline{f}=\int_\Omega fdx$$

次元に依存しないソボレフ不等式

T.S. Chemotaxis, Reaction, Network, World Scientific, Singapore, 2018 第4章

推測統計のまとめ

・データ分布は標準確率空間上の確率変数と見なすことができる
・この確率変数の平均、分散は1変量データとしての平均、分散（統計量）と一致する
・このとき標本抽出は独立同分布に従う確率変数の列として定式化される

・不偏性、一致性、一様最小分布性は統計的推定の基準である
・最尤推定はデータの独立性の仮定のもとに観測データが最も起こりやすいパラメータを選ぶ手法である
・情報量統計は母集団‐標本の図式に依存せず観測データに即してモデル分布を推論する方法である
・KL情報量は統計力学の手法により導出され幾何学や解析学でも使われる指標である
・AICは観測データからモデルとパラメータを同時に評価する基準である

6. 孤立系の均質化

点渦系の平衡統計力学

2D Euler Equation
(simply connected domain)

$$v_t + (v \cdot \nabla)v = -\nabla p$$
$$\nabla \cdot v = 0$$
$$\nu \cdot v|_{\partial\Omega} = 0$$

$$\nabla = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \end{pmatrix}$$ gradient

vorticity

$$\omega = \nabla^\perp v$$
$$\nabla^\perp = \begin{pmatrix} \frac{\partial}{\partial x_2} \\ -\frac{\partial}{\partial x_1} \end{pmatrix}$$
$$x = (x_1, x_2)$$

vortex equation

$$\omega_t + v \cdot \nabla\omega = 0$$
$$v = \nabla^\perp \psi$$
$$\Delta\psi = -\omega \quad \text{stream function}$$
$$\nabla^\perp \cdot \nabla^\perp = \Delta \qquad \psi|_{\partial\Omega} = 0$$

point vortices

$$\omega(dx, t) = \sum_{i=1}^{N} \alpha_i \delta_{x_i(t)}(dx)$$

Kirchhoff equation

$$\alpha_i \frac{dx_i}{dt} = \nabla_i^\perp H, \ 1 \le i \le N$$

Hamiltonian

$$H = \sum_i \frac{\alpha_i^2}{2} R(x_j) + \sum_{i<j} \alpha_i \alpha_j G(x_i, x_j)$$

Green's function

$$-\Delta G(x, x') = \delta_{x'}(dx)$$
$$G(x, x')|_{\partial\Omega} = 0$$
$$(x, x') \in \overline{\Omega} \times \Omega$$

Robin function

$$R(x) = \left[G(x, x') + \frac{1}{2\pi} \log|x - x'| \right]_{x'=x}$$

Onsager 49

Hamiltonian $H = \sum_i \frac{\alpha_i^2}{2} R(x_j) + \sum_{i<j} \alpha_i \alpha_j G(x_i, x_j)$

$H = \hat{H}_N(x_1, \cdots, x_N) \qquad N \gg 1$ total energy

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i}, \quad \frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i}, \quad 1 \leq i \leq N$$

$$p_i = p_i(t), \; q_i = q_i(t) \; \in \mathbf{R}^2$$

micro-canonical ensemble $\quad \mathbf{R}^{4N}/\{H = E\}$

$$x = (q_1, \ldots, q_N, p_1, \ldots, p_N)$$

co-area formula $\quad dx = dE \cdot \frac{d\Sigma(E)}{|\nabla H|}$

$$d\Sigma(E) \; \leftrightarrow \; \{x \in \mathbf{R}^{4N} \mid H(x) = E\}$$

$\Longrightarrow$ canonical ensemble

thermal equilibrium

Gibbs measure $\quad \mu^{E,N} = \frac{1}{W(E)} \cdot \frac{d\Sigma(E)}{|\nabla H|}$

weight factor $\quad W(E) = \int_{H=E} \frac{d\Sigma(E)}{|\nabla H|}$

inverse temperature $\quad \beta = \frac{\partial}{\partial E} \log W(E) = \frac{\Theta''(E)}{\Theta'(E)}$

$$\Theta(E) = \int_{H<E} dx = \int_{-\infty}^{E} W(E')dE'$$

bounded monotone

$E \gg 1 \Rightarrow \beta < 0$ ordered structure in negative temperature

Joyce-Montgomery 73

$\boxed{\text{micro-canonical statistics}}$

$$\mathbf{R}^{4N}/\{H=E\}$$

$$x=(q_1,\dots,q_N,p_1,\dots,p_N)$$

$$dx = dE\cdot\frac{d\Sigma(E)}{|\nabla H|}$$

$$d\Sigma(E)\ \leftrightarrow\ \{x\in\mathbf{R}^{4N}\mid H(x)=E\}$$

micro-canonical measure
$$d\mu^{E,N}=\frac{1}{W(E)}\cdot\frac{d\Sigma(E)}{|\nabla H|}$$

weight factor
$$W(E)=\int_{\{H=E\}}\frac{d\Sigma(E)}{|\nabla H|}$$

$\boxed{\text{canonical statistics}}$

$$\mathbf{R}^{4N}/\{T\}$$

inverse temperature
$$\beta=1/(kT)$$

canonical measure
$$d\mu^{\beta,N}=\frac{e^{-\beta H}dx}{Z(\beta,N)}$$

weight factor
$$Z(\beta,N)=\int_{\mathbf{R}^{4N}}e^{-\beta H}dx$$

thermo-dynamical relation
$$\beta=\frac{\partial}{\partial E}\log W(E)$$

micro-canonical probability measure
$$\mu^n=\mu^n(dx_1,\dots,dx_n)$$

one point pdf
$$\rho_1^n(x_i)dx_i = \int_{\Omega^{n-1}}\mu^n(dx_1\dots dx_{i-1}dx_{i+1}dx_n)$$

equal a priori probability (independent of i)

k-point reduced pdf
$$\rho_k^n(x_1,\dots,x_k)dx_1\dots dx_k = \int_{\Omega^{n-k}}\mu^n(dx_{k+1},\dots,dx_n)$$

stationary point vortices
$$\omega_N(x)dx=\sum_{i=1}^N\alpha\delta_{x_i}(dx)$$

$\Longrightarrow$

$$\begin{aligned}\langle\omega_N(x)\rangle &= \sum_{i=1}^N\int_{\Omega^N}\alpha\delta(x_i-x)\mu^N(dx_1\dots dx_N)\\ &= N\alpha\rho_1^N(x)\end{aligned}$$

phase mean

high energy limit
(single intensity)

$$\alpha_i = \hat{\alpha},\ N \uparrow +\infty,\ \hat{\alpha} N = 1$$
$$\hat{H}_N = H,\ \hat{\alpha}^2 N \hat{\beta} = \beta$$

$\Longrightarrow$ two point pdf compatibility

$$\hat{H}_N(x_1, \ldots, x_N) = \sum_i \frac{\alpha_i^2}{2} R(x_j) + \sum_{i<j} \alpha_i \alpha_j G(x_i, x_j)$$

Boltzmann
$$\rho = \frac{e^{-\beta\psi}}{\int_\Omega e^{-\beta\psi}}$$

duality $\updownarrow$

Poisson
$$\psi = \int_\Omega G(\cdot, x')\rho(x')dx'$$

energy
$$\tilde{E} = H$$

inverse temperature
$$\tilde{\beta} = \frac{\partial}{\partial \tilde{E}} \log W(\tilde{E})$$

weight factor
$$W(\tilde{E}) = \int_{H=\tilde{E}} \frac{d\Sigma_{\tilde{E}}}{|\nabla H|}$$

mean field limit
$$\lim_{N\to\infty} \langle \omega_N(x) \rangle = \rho(x) = \lim_{N\to\infty} N\alpha\rho_1^N(x)$$

propagation of chaos
(factorization property)
$$\rho_k^N \rightharpoonup \rho^{\otimes k} = \Pi_{i=1}^k \rho(x_i)$$

(Suzuki's uniqueness theorem)

<u>rigorous derivation</u>

Caglioti-Lions-Marchioro-Pulvirenti 92, 95. Kiessling 93

1. Bounded Boltzmann weight factors {z}
2. Uniqueness of the solution to the limit equation

$\Longrightarrow$

1. convergence to the limit
2. canonical-micro canonical equivalence in the limit
3. propagation of chaos

OK if $\beta > -8\pi$

スモルコフスキー・ポアソン方程式

$u = u(x,t) \geq 0$ density

total mass conservation $\quad \frac{d}{dt}\|u(t)\|_1 = 0$

$$u_t = \nabla \cdot (\nabla u - u\nabla v), \ -\Delta v = u, \quad \left(\frac{\partial u}{\partial \nu} - u\frac{\partial v}{\partial \nu}, v\right)\Big|_{\partial\Omega} = 0$$

$G(x,x') = G(x',x)$ Green's function

$u \otimes u = u(x,t)u(x',t)\ dxdx'$

Helmholtz

free energy decreasing
$$\frac{d}{dt}\left\{\int_\Omega u(\log u - 1) - \frac{1}{2}\int\int_{\Omega\times\Omega} G(x,x')u\otimes u\right\} = -\int_\Omega u|\nabla(\log u - v)|^2 \leq 0$$

Boltzmann-Shannon entropy
$$H_1(u) = \int_\Omega u(\log u - 1)$$

Non-local Rao entropy
$$H_2(u) = \frac{1}{2}\iint_{\Omega\times\Omega} G(x,x')u\otimes u$$

Relative Boltzmann-Shannon entropy
$$H_1(u \mid \overline{u}) = \int_\Omega u\log\left(\frac{u}{\overline{u}}\right) - (u - \overline{u})\ dx$$

$$\frac{dH_1}{dt} = \int_\Omega -4|\nabla u^{\frac{1}{2}}|^2 + u^2\ dx$$

Relative non-local Rao entropy
$$H_2(u \mid \overline{u}) = \frac{1}{2}\iint_{\Omega\times\Omega} G(x,x')(u-\overline{u})\otimes(u-\overline{u})$$

$$\frac{dH_2}{dt} = \int_\Omega -u^2 + u|\nabla v|^2\ dx$$

$$\Longrightarrow \quad \frac{d}{dt}(H_1 - H_2) = -\int_\Omega u|\nabla(\log u - v)|^2$$

$$H_1(u \mid \overline{u}) = H_1(u) - H_1(\overline{u}), \ \overline{u} \in \mathbf{R}_+$$

散逸反応拡散系

$$\tau_j \frac{\partial u_j}{\partial t} - d_j \Delta u_j = f_j(u) \ \text{in} \ Q_T$$

$$\left. \frac{\partial u_j}{\partial \nu} \right|_{\partial\Omega} = 0, \ \ u_j|_{t=0} = u_{j0}(x)$$

$\Omega \subset \mathbf{R}^n$ bounded domain, $\partial\Omega$ smooth

$Q_T = \Omega \times (0,T) \qquad 1 \leq j \leq N$

ν outer unit normal

$\tau = (\tau_j) > 0, \ \ d = (d_j) > 0$

$u_0 = (u_{j0}) \geq 0$ smooth

[local. Lipschitz cont.]

$f_j : \mathbf{R}^N \to \mathbf{R}, \ \ 1 \leq j \leq N$

loc. Lipschitz cont.

$\Longrightarrow$

$\exists 1$ classical solution local-in-time

$T \in (0, +\infty]$ maximal existence time

[quasi-positive]

$$f_j(u_1, \cdots, u_{j-1}, 0, u_{j+1}, \cdots, u_n) \geq 0, \ \ \forall j$$

$$0 \leq u_0 = (u_{j0}) \in \mathbf{R}^N \Longrightarrow u = (u_j(\cdot, t)) \geq 0$$

[mass dissipation]

$$\sum_{j=1}^N f_j(u) \leq 0, \ \ u = (u_j) \geq 0$$

$$\Longrightarrow \frac{\partial}{\partial t}(\tau \cdot u) - \Delta(d \cdot u) \leq 0$$

$$\left. \frac{\partial u}{\partial \nu} \right|_{\partial\Omega} = 0$$

$$\|\tau \cdot u(t)\|_1 \leq \|\tau \cdot u_0\|_1$$

[quadratic]

$$|\nabla f_j(u)| \leq C(1 + |u|), \ \ \forall j$$

Theorem (Fellner-Morgan-Tang 20, 21)

$$T = +\infty \qquad \|u(\cdot, t)\|_\infty \leq C$$

Examples

chemical reaction $A_1 + \cdots + A_m \rightleftharpoons A_{m+1} + \cdots + A_N$

$$\tau_j \frac{\partial u_j}{\partial t} - d_j \Delta u_j = \chi_j f(u), \ \left. \frac{\partial u_j}{\partial \nu} \right|_{\partial\Omega} = 0$$

micro-canonical ensemble

$$f(u) = \prod_{j=1}^{m} u_j - \prod_{j=m+1}^{N} u_j, \ \chi_j = \begin{cases} -1, & 1 \leq j \leq m \\ 1, & m+1 \leq j \leq N \end{cases}$$

spatially homogeneous stationary state $\overline{w} = \frac{1}{|\Omega|} \int_\Omega w$

$$0 \leq \exists 1 z = (z_j) \in \mathbf{R}^N, \ f(z) = 0$$

$$z_i + z_k = \overline{u}_{i0} + \overline{u}_{k0}, \ 1 \leq i \leq m, \ m+1 \leq k \leq N$$

$\Longrightarrow$ $z = (z_j) > 0$

Theorem $m = 2, \ N = 4$ (quadratic)

$\Longrightarrow$ $T = +\infty$ $\|u(\cdot, t) - z\|_\infty \leq Ce^{-\delta t}$

Devillettes-Fellner 08, Pierre-S.-Zou 17, Pierre-S.-Umakoshi 18

$$\Phi(s) = s(\log s - 1) + 1 \geq 0$$

relative entropy (diversity) entropy

$$E(w \mid v) = \int_\Omega v \ \Phi\left(\frac{w}{v}\right), \ E(w) = \int_\Omega \Phi(w)$$

$$E(u) = \sum_{j=1}^{N} \tau_j E(u_j), \ E(u \mid z) = \sum_{j=1}^{N} \tau_j E(u_j \mid z_j)$$

$\Longrightarrow$ $E(u|z) = E(u) - E(z)$

$$\boxed{\frac{d}{dt} E(u) = -D(u)} \qquad D(u) = 4 \sum_{j=1}^{N} d_j \|\nabla \sqrt{u_j}\|_2^2 + \int_\Omega f(u) \log \frac{\prod_{j=m+1}^{N} u_j}{\prod_{j=1}^{m} u_j}$$

[logarithmic Sobolev] $D(u) \geq 2\delta E(u|z)$

[Csiz'ar-Kullback] $\|v - \overline{v}\|_1^2 \leq 4\overline{v} E(v|\overline{v})$

気体粒子の非平衡統計力学　ブラウン運動を導入して粒子運動（小正準集団）を平均化 ⟶ 熱力学法則を実現

質量保存
エネルギー保存
エントロピー増大

点渦系　Chavanis（08）→ S.（14）→ Sawada-S.（17）

緩和時間モデル（非局所項あり）　数学解析　パッチモデルとの整合性

定常解の構造
非定常解の挙動

格子モデル ⟶ 連続極限

非局所項のない孤立系

Streater（97a）

粒子密度 $f = f(x,t)$, 温度 $\theta = \theta(x,t)$　熱ポテンシャル $v = v(x)$

古典粒子モデル　$f_t = f_{xx} + (\frac{v_x f}{\theta})_x, \ \theta_t = \theta_{xx} + v_x(\frac{v_x f}{\theta} + f_x)$

Smoluchowski 方程式　状態方程式＋拡散項

量子粒子モデル　$f_t = (\frac{\theta}{v_x} f_x)_x + f_x, \ \theta_t = \theta_{xx} + \theta f_x + v_x f$

Streater（97b）　古典粒子モデルを高次元化、全空間での局所適切性

S. (23)　古典粒子モデルを有界領域で定式化、エントロピー生成の有界性のもとで空間2次元までの時間大域解の存在

S. (24)　定常状態の一意性、量子粒子モデルの高次元化と時間大域解の存在（十分条件）

古典粒子　　$\Omega \subset \mathbf{R}^n$ 有界　$\partial\Omega$ 滑らか　$v = v(x)$ 滑らか　ν 外向き単位法ベクトル

$$f_t = \Delta f + \nabla \cdot (\frac{f}{\theta}\nabla v),\ \theta_t = \Delta\theta + \nabla v \cdot (\frac{f}{\theta}\nabla v + \nabla f) \quad \text{in } \Omega \times (0,T) \qquad T = T_{\max} \in (0,\infty]$$

$$(\frac{\partial f}{\partial \nu} + \frac{f}{\theta}\frac{\partial v}{\partial \nu}, \frac{\partial \theta}{\partial \nu})\Big|_{\partial\Omega} = 0, \quad (f,\theta)|_{t=0} = (f_0(x), \theta_0(x)) > 0$$ 滑らか　⟹　時間局所古典解の一意存在

正値性保存　$f(x,t), \overset{\text{非自明}}{\theta(x,t)} > 0$　質量保存　$M = \int_\Omega f\ dx$　エネルギー保存　$E = \int_\Omega fv + \theta\ dx$

エントロピー増大　$\overset{\text{エントロピー}}{S} = \int_\Omega -f(\log f - 1) + \log\theta\ dx$　$\overset{\text{エントロピー生成}}{\frac{dS}{dt}} = \int_\Omega |\nabla \log\theta|^2 + \frac{1}{f}|b|^2\ dx \geq 0$　$b = \nabla f + \frac{f}{\theta}\nabla v$

定常解　$M = \int_\Omega f\ dx,\ E = \int_\Omega fv + \theta\ dx,\ \frac{dS}{dt} = \int_\Omega |\nabla \log\theta|^2 + f\left|\nabla \log f + \frac{\nabla v}{\theta}\right|^2 dx = 0$

⟹ $\nabla\theta = 0,\ \nabla(\log f + \frac{v}{\theta}) = 0$ ⟹ $0 < \theta \in \mathbf{R}$　$\log f = -\frac{v}{\theta} + c,\ c \in \mathbf{R}$ ⟹ $f = \frac{Me^{-v/\theta}}{\int_\Omega e^{-v/\theta}dx}$

Boltzmannの式

$$E = |\Omega|\theta + \frac{M}{\int_\Omega e^{-v/\theta}dx}\int_\Omega e^{-v/\theta}v\ dx$$

$$\boxed{\frac{\int_\Omega e^{sv}v\ dx}{\int_\Omega e^{sv}} = \frac{E}{M} + \frac{|\Omega|}{M}s^{-1}}$$

$s = -\theta^{-1} < 0$

定理 (S. 23) $\quad n \leq 2, \quad \left.\frac{\partial v}{\partial \nu}\right|_{\partial\Omega} = 0 \quad \limsup_{t \uparrow T} \frac{dS}{dt} < +\infty \;\Rightarrow\; T = +\infty, \; \|f(\cdot,t), \theta(\cdot,t), \theta(\cdot,t)^{-1}\|_{C^{2+\alpha}} \leq C \qquad (1)$

定理 1 $\quad v \notin \mathbf{R} \;\longrightarrow\; \forall \frac{E}{M} \in \mathbf{R}, \; \exists 1 s < 0 \qquad \frac{\int_\Omega e^{sv} v \, dx}{\int_\Omega e^{sv}} = \frac{E}{M} + \frac{|\Omega|}{M} s^{-1}$

放物型強最大原理（ω極限集合元の正値性）

系 $\quad v \notin \mathbf{R} \quad (1) \;\longrightarrow\; \exists 1 \theta_* > 0, \; \lim_{t \uparrow +\infty} \|f(\cdot,t) - f_*, \theta(\cdot,t) - \theta_*\|_{C^2} = 0, \; f_* = \frac{M e^{-v/\theta_*}}{\int_\Omega e^{-v/\theta_*} dx}$

$v \notin \mathbf{R}$

証明 $\quad I(s) = \frac{d}{ds} \log \left(\int_\Omega e^{sv} dx \right) = \frac{\int_\Omega e^{sv} v dx}{\int_\Omega e^{sv} dx} \;\longrightarrow\; \frac{dI}{ds} = \int_\Omega v^2 \rho_s dx - \left(\int_\Omega v \rho_s dx \right)^2 > 0, \; \rho_s = \frac{e^{sv}}{\int_\Omega e^{sv} dx}$

一意性

補題 $\quad w \in C(\overline{\Omega}), \; \min_{\overline{\Omega}} w = 0 \;\longrightarrow\; \lim_{s \downarrow -\infty} \int_\Omega \frac{e^{sw}}{\int_\Omega e^{sw} dx} w dx = 0$

存在

$\Rightarrow \quad \lim_{s \downarrow -\infty} \int_\Omega \frac{e^{sv}}{\int_\Omega e^{sv} dx} v dx = v_{\min} \leq \int_\Omega \frac{f_0}{\int_\Omega f_0 dx} v dx < \frac{1}{\int_\Omega f_0 dx} \int_\Omega f_0 v + \theta_0 \; dx = \frac{E}{M}$

量子粒子モデル $f_t = (\frac{\theta}{v_x} f_x)_x + f_x,\ \theta_t = \theta_{xx} + \theta f_x + v_x f$ $\Longrightarrow$ 高次元化

$\Omega \subset \mathbf{R}^n$ 有界 $\partial\Omega$ 滑らか ν 外向き単位法ベクトル $v(x)$ 滑らか $b = |\nabla v| > 0 \text{ on } \overline{\Omega}$ $\omega \stackrel{\text{def}}{=} \frac{\nabla v}{b}$

$$f_t = \nabla \cdot (\frac{\theta}{b} \nabla f) + \nabla \cdot (\omega f),\ \theta_t = \Delta\theta + \theta\nabla \cdot (\omega f) + bf \quad \text{in } \Omega \times (0, T)$$

$$(\frac{\theta}{b}\frac{\partial f}{\partial \nu} + \nu\omega f, \frac{\partial \theta}{\partial \nu})\Big|_{\partial\Omega} = 0, \quad (f, \theta)|_{t=0} = (f_0(x), \theta_0(x)) > 0$$

$n = 1,\ b = v_x > 0,\ \omega = 1$

正値性保存 $f(x,t), \theta(x,t) > 0$

質量保存 $M = \int_\Omega f\ dx$

エネルギー保存 $E = \int_\Omega fv + \theta\ dx$

エントロピー増大 $S = \int_\Omega -f(\log f - 1) + \log\theta\ dx$

$$\frac{dS}{dt} = \int_\Omega |\nabla \log\theta|^2 + \frac{\theta f}{b}\left|\nabla \log f + \frac{\nabla v}{\theta}\right|^2\ dx$$

定常状態は古典粒子モデルと一致、一意

定常状態 $\theta_* > 0,\ f_* = \frac{Me^{-v/\theta_*}}{\int_\Omega e^{-v/\theta_*} dx}$

$$\frac{\int_\Omega e^{sv} v\ dx}{\int_\Omega e^{sv}} = \frac{E}{M} + \frac{|\Omega|}{M} s^{-1} \qquad s = -\theta_*^{-1} < 0$$

定理２ $n \le 2,\ (\omega\cdot\nu)f|_{\partial\Omega} \ge -C,\ \dfrac{dS}{dt} \le C,\ 0 \le t < T$ ⟹ $\|f(\cdot,t)^{1/2}\|_{H^1} + \|\theta(\cdot,t)^{-1}\|_\infty \le C,\ 0 \le t < T$

$\|\theta\|_{C^\mu(\Omega\times(0,T))} \le C,\ 0 < \exists\mu < 1$ (2)

定理３ (2) $T = +\infty$

$t_k \to +\infty,\ \lim_{k\to\infty} \dfrac{dS}{dt}(t_k) = 0$ ⟹ $\lim_{k\to\infty} \|\theta(\cdot,t_k) - \theta_*\|_\infty = 0,\ f(\cdot,t_k)^{1/2} \rightharpoonup f_*^{1/2} \text{ in } H^1(\Omega)$

定理４ $n = 1,\ T = +\infty$

$(\omega\cdot\nu)f|_{\partial\Omega} \ge -C,\ \dfrac{dS}{dt} \le C,\ \limsup_{t\uparrow+\infty} S(t) < +\infty$ ⟹ $\lim_{t\uparrow+\infty} \|\theta(\cdot,t) - \theta_*, f(\cdot,t) - f_*\|_\infty = 0$

原理 孤立系（小正準集団）は空間均質化する　閉じた系（正準集団）は定常状態の集合が支配する
（非線形スペクトル力学、自己組織化のポテンシャル）

注意 Trudinger-Moser 不等式
Chang-Yang 88. Proposition 2.3

Harnack不等式 ⟹ 非負解の強最大原理（楕円型）
Gilbarg-Trudinger 83. Theorem 8.20

空間均質化解明の３つの方法

1．双対変分原理
2．相対エントロピー
3．勾配不等式

$$f_t = \nabla\cdot\left(\frac{\theta}{b}\nabla f\right) + \nabla\cdot(\omega f),\ \theta_t = \Delta\theta + \theta\nabla\cdot(\omega f) + bf$$

S. Mathematical study of Streater’s model on Brownian particle gas, DIE, accepted for publication (2024)