

科目名	データ駆動科学C
講義題目(テーマ)	データ駆動科学を用いた物質科学
担当教員	東京大学 福島孝治
年度・学期	2023年 集中
単位数	1

学修成果とその割合	
1.高度な専門的知識・技能及び研究力	70
2.学際的領域を理解できる深奥な教養力	30
3.グローバルな視野と行動力	0
4.地域社会を牽引するリーダー力	0
その他	0

使用言語	「日本語」による授業
教科書・資料の言語	「日本語」のテキスト
実務経験を活かした授業	非該当
授業の形態	講義
対面・遠隔の別	遠隔形式
授業の方法	Zoomを用いた遠隔授業と、オンデマンド受講

授業の目的	マルコフ連鎖モンテカルロ法によるベイズ推定を用いた物質科学について理解する
授業の概要	物質科学研究には実験・計測は欠かせない。その実験・計測によって得られるデータの解析もまた研究活動には不可欠である。その解析方法はいつ学ぶのかというと、多くの場合は学生実験でのデータ解析であるが、その後、大学院において研究室でのOn the Job Training(OJT)的な学びになり、近年はそこで機械学習に触れることになるだろう。学部教育での普遍的な統計学と昨今話題の機械学習との間には大きなギャップがあるように見える。本講義では、ベイズ統計の枠組みからデータ科学の手法を概観し、学生実験のときの最小二乗法と研究の現場で行なうベイズ的データ解析の溝を埋め、ベイズ統計の実践的な方法論であるマルコフ連鎖モンテカルロ法の解説を行なう。モンテカルロ法は乱数を使う数値計算手法の総称であるが、特に確率分布からのサンプリング技法としての側面に注目し、その数理から先端的な方法を取り上げる。

学修目標	
A水準 (到達すれば「優」に相当)	(1). ベイズ統計の基礎、(2).スパースモデリングの数理、(3).マルコフ連鎖モンテカルロ法の原理、(4).メトロポリス法、(5). ハミルトンモンテカルロ法、(6).拡張アンサンブル法、(7).高次元モンテカルロ積分、(8).ポピュレーションモンテカルロ法について、十分に理解し、講義内容を他人に正確に説明できる。

C水準（到達すれば「可」に相当）	(1). ベイズ統計の基礎、(2).スパースモデリングの数理、(3).マルコフ連鎖モンテカルロ法の原理、(4).メトロポリス法、(5). ハミルトンモンテカルロ法、(6).拡張アンサンブル法、(7).高次元モンテカルロ積分、(8).ポピュレーションモンテカルロ法について、概ね理解し、講義内容の要点をまとめることができる。
評価方法・基準	Moodleで提出されたレポートの到達度(100%) から評価する。

各回の授業内容		
回	授業テーマ (5文字以上100文字以内)	内容概略(10文字以上200文字以内)
1	ベイズ統計の基礎	ベイズ統計を概観し、最適化からサンプリングへの移行を通じて、推定の不確実度の評価の意義を学ぶ。
2	マルコフ連鎖モンテカルロ法の基礎1	メトロポリス法からギブスサンプリングまでを実践的に学ぶ
3	マルコフ連鎖モンテカルロ法の基礎2	ハミルトンモンテカルロ法を実践的に学び、その有効性と限界を認識する
4	マルコフ連鎖モンテカルロ法の数理	マルコフ連鎖の数理を学び、モンテカルロ法の原理を理解する
5	マルコフ連鎖モンテカルロ法の実践と先端的手法1	拡張アンサンブル法の原理を学び、応用例を概観する
6	マルコフ連鎖モンテカルロ法の実践と先端的手法2	ベイズ統計における周辺尤度計算など実践的な計算例を学ぶ
7	マルコフ連鎖モンテカルロ法の実践と先端的手法3	近年の機械学習の方法とモンテカルロ法の統合的応用について概観する
8	ポピュレーションモンテカルロ法の数理と応用	並列計算向きのモンテカルロ法であるポピュレーションモンテカルロ法の数理と応用を学ぶ

授業外学修時間の目安	本科目は、45時間の学修が必要な内容で構成されている。授業は16時間分（2h×8コマ）となるため、29時間分相当の事前・事後学修（課題等含む）が、授業の理解を深めるために必要となる。
------------	---

キーワード	マルコフ連鎖モンテカルロ法、ポピュレーションモンテカルロ法
テキスト	授業の際に資料を配布する。
参考文献	

オフィス・アワー	データ駆動型社会を担う人材育成プログラム事務室に連絡をとること
担当教員への連絡方法	データ駆動型社会を担う人材育成プログラム事務室に連絡をとること
担当教員からのメッセージ	学生の頃、実験データの誤差解析など退屈で仕方がなかった。何も主張する気がなかったから、ある意味で当然である。ところが、大学院生になって研究がはじまり、自分の主張や議論がこのデータ解析に基づくようになると、どのくらい自信を持ってよいのか不安になる。誤差は自信の根拠になることによりやく気づくようになるが、今度は誤差の求め方がわからない。解析が複雑になればなるほど、誤差伝搬などは窮屈になる。マルコフ連鎖モンテカルロ法は自由なデータ解析を可能にする重要なツールである。データ解析で誤差棒が必要な方はぜひ議論しましょう。