科目ナンバリング	U-LAS11 10007 LJ55						
授業科目名<英訳>	データ分析基礎 Basic Data Analysis			担当者所属職名・氏名	情報学研究科 教授　林　和則		
群	自然科学科目群	分野(分類)	データ科学(基礎)			使用言語	日本語
旧群	B群	単位数	2単位	週コマ数	1コマ	授業形態	講義（対面授業科目）
開講年度・開講期	2026・前期	曜時限	木4	配当学年	全回生	対象学生	全学向

[授業の概要・目的]

本授業の目的は、データ分析の基本的な事項を学び、実際に大規模データ解析を行うための基礎的な技術を身につけることである。大規模データ解析や最新のデータ分析手法の適用は、既存のソフトウェアを用いて簡単に実行できるとは限らない。そこで、本授業では、データ分析の基礎となる線形代数についても講義する。ただし、厳密な数学的証明は必要最小限に抑え、直観的な理解を深めることを目標とする。本講義は、文部科学省のモデルカリキュラム（応用基礎レベル）の、データサイエンス基礎・データエンジニアリング基礎・AI基礎をバランスよくカバーする形で構成されている。

はじめに、データを扱う上で必要不可欠な確率変数・平均・分散などの統計の基礎的な概念について解説する。また、仮説検定の考え方を述べ、いくつかの例を用いて解説する。その後、複数の変数の関係性を調べる多変量解析を行う際に必要となる複素数と線形代数の重要事項について講義する。次に、多変量解析の中心的な解析対象となる相関行列の性質について解説し、複素数で表される信号やシステムを扱う際に有用なウィルティンガー微分について説明する。そして，多くのデータ分析の場面で登場する線形観測モデルの逆問題の解き方について講義する。さらに、劣決定の線形観測モデルにおいて、未知ベクトルのスパース性を利用して信号の推定を行う手法である圧縮センシングの基本的な考え方やアルゴリズムについて解説する。最後に、最近の機械学習に基づく人工知能(AI)の基礎と展望について講義する。

本講義の単位（2単位）を修得することで、文部科学省が定める数理・データサイエンス・AI教育プログラム応用基礎レベル(MDASH Advanced Literacy)修了証の取得が可能である。修了証取得の手続きについては、講義内で担当教員より指示がある。

[到達目標]

1. 確率変数・平均・分散など統計の基礎的な内容、および、仮説検定について理解する。
2. 多変量解析の基礎である相関行列の性質や線形逆問題についての概念を理解し、データ分析に応用できるようになる。
3. 統計ソフトRを使いこなす知識を身につけ、実際に簡単なデータ分析を行えるようになる。
4. データ分析に登場する最適化問題とそれを解くアルゴリズムについて、基本的な考え方を理解する。

[授業計画と内容]

授業回数はフィードバックを含め全15回とする。
開講にあたっては、受講生の所属するキャンパスの配置や受講形態にも配慮し、一部メディア授業を取り入れることがある。

・ガイダンス（1回）
授業の概要について説明する。その後、データ駆動型社会、Society 5.0、ビッグデータ、プライバシー保護、個人情報の取り扱いについても解説する。

・確率・統計の基礎（2回）

データ分析基礎(2)へ続く↓↓↓

データ分析基礎(2)

確率変数・平均・分散など統計の基礎的な内容について講義する。

・仮説検定（2回）
仮説検定の基本的な考え方を講義する。

・複素数・線形代数の基礎（2回）
複素数の定義から始め，その性質や演算規則について解説する。
さらに，実および複素行列・ベクトルの性質や，行列式，逆行列，固有値分解，特異値分解について解説する。

・統計ソフトR（1回）
統計ソフトRの基本的な使い方について，実例を交えながら解説する。

・線形逆問題の基礎（3回）
様々なデータ分析において重要な役割を果たす相関行列の性質について説明し、勾配法で利用されるウィルティンガー微分について解説する。さらに、線形観測モデルを導入し、その逆問題を解くための基本的な手法について詳細に説明する。

・圧縮センシングの基礎（2回）
劣決定の線形観測モデルにおいて、未知ベクトルのスパース性を利用して信号の推定を行う手法である圧縮センシングの基本的な考え方やアルゴリズムについて解説する。

・AI・機械学習の基礎と展望（1回）
AIの歴史や背景からはじめて、機械学習（教師あり学習・教師なし学習・強化学習）の考え方について説明する。ニューラルネットワークの原理や深層学習、生成AIなどについても概説する。

・フィードバック（1回）

[履修要件]

全学共通科目の統計入門、あるいは数理統計で学習した知識を一部前提とするので、これらのいずれかを受講済であることを推奨する。

[成績評価の方法・観点]

定期試験、小レポート（平常点）により総合的に評価する。

・定期試験（配点50点）
講義で解説したデータ分析の基本的な原理や理論を理解できているかを評価する。

・小レポート（平常点）（配点50点）
講義で解説したデータ分析手法などをプログラミングで演習する小レポート課題を２回程度課す。

[教科書]

教科書は使用しない。
資料が必要な場合には、授業中に配布する。

[参考書等]

（参考書）
京都大学　データ科学イノベーション教育研究センター『講義実録　統計入門』（現代図書, 2023）

データ分析基礎(3)へ続く↓↓↓

データ分析基礎(3)

ISBN:978-4-434-31857-3

[授業外学修（予習・復習）等]

行列やベクトルについては授業中にも解説するが、行列やベクトルの扱いに慣れていない場合は予習あるいは復習をすることが望ましい。
講義中に説明したアルゴリズムを、何らかのプログラミング言語で実装して動かしてみることを推奨する。

[その他（オフィスアワー等）]

授業時間外で質問がある場合には、下記のアドレスにメールで連絡すること。
林　和則（はやしかずのり）hayashi.kazunori.4w@kyoto-u.ac.jp

[主要授業科目（学部・学科名）]