

平成25年度研究科横断型教育プログラム（Aタイプ）授業科目

開講方式	Aタイプ (研究科 開講型)	研究科名	情報学研究科 (附属情報教育 推進センター)	カテゴリー	情報活用・計算科 科学科目群	横断 区分	理系横断型		
授業科目名 (英訳)	ビッグデータの計算科学 (Computational Science for Big Data)			講義担当者 所属・氏名	情報学研究科・木村欣司 情報学研究科・關戸啓人 高等教育研究開発推進機 構・小山田耕二 情報学研究科・中村佳正	開講 場所	吉田南 学術情報メデ ィアセンター南館 204		
配当 学年	修士 博士後期 専門職	単位 数	2単位	開講 期	後期	曜 時 限	水5限 (16:30-18:00)	授業 形態	講義 (演習を含む)
〔授業の概要・目的〕									
<p>近年のコンピュータの進歩や情報基盤技術の整備に伴って、クラウドコンピューティングなどのインターネットを介して行われる社会活動から生成されるデータの量、あるいは、計算科学の重要な技法であるコンピュータシミュレーションを通じて得られるデータの量は、日々増加の一途をたどっている。それらのビッグデータを分析、可視化するための手法を学ぶことが、この科目の目的である。特に、C言語を利用して、大次元の疎行列に対するデータ分析の演習を行う。</p> <p>大次元疎行列(大次元隣接行列)は、重み付き有向グラフを表現する能力を持っている。よって、大規模データ(ビッグデータ)を記述する際に、多様な分析対象を表現することが可能である。その行列の特徴量、すなわち、分析対象の特徴量を抽出する際に、最も一般的でかつ普遍的な手法は、特異値分解を行うことである。それ以外にも、特異値分解は、最小2乗法、主成分分析、独立成分分析といった解析したいデータがはじめてから行列の形式で表現されている問題への幅広い応用も可能である。よって、本科目は、受講者が特異値分解をおこなうプログラムをソースコードのレベルから作成することにより、大規模データを分析するための基本的な技術を習得することを目的とする。ソースコードのレベルからプログラムを作成することは、プログラミング技術を習得することにもつながる。本科目では、C言語の基本文法などの基礎的な話題から演習を開始する。よって、過去にC言語を学んだことのない学生の受講も歓迎する。</p>									
【研究科横断型教育の概要・目的】									
<p>「ビッグデータの可視化」、「密行列の特異値分解」、「疎行列の特異値分解」を通じて、大規模データ(ビッグデータ)から重要な情報を取り出す分析能力、可視化する能力を身につけることを目指す科目である。大規模データ(ビッグデータ)を分析する能力は、研究分野を問わず、重要である。</p>									
〔授業計画と内容〕									
<p>○ガイダンス(木村欣司/1回 講義)</p> <p>計算科学は、数学的モデルとその定量的評価法を構築し、計算機を駆使して科学技術上の問題を解決する学問分野である。計算科学概論、計算科学の応用について講述する。</p> <p>○クラウドコンピューティング入門(木村欣司/1回 講義)</p> <p>クラウドコンピューティングの基本的な話題について解説を行う。</p> <p>○ビッグデータの可視化(小山田耕二/3回 講義)</p> <p>ビッグデータを視覚的に理解するための技法について解説する。</p> <p>○密行列の特異値分解法(中村佳正/3回 講義)</p> <ol style="list-style-type: none"> (1) 大次元疎行列(大次元隣接行列)と重み付き有向グラフの関係についての解説 (2) 重み付き有向グラフを用いて表現できるデータの例を列挙 (3) データを分析するための統計的手法についての解説 <ol style="list-style-type: none"> I. 最小2乗法 II. 主成分分析 III. 独立成分分析 (4) ビッグデータの特徴量の抽出における特異値分解の重要性についての解説 (5) 大次元疎行列の特異値分解の準備として、中規模サイズの密行列の特異値分解の具体的な方法についての解説 (6) 特異値分解における、ハウスホルダー変換による前処理と後処理の解説 (7) QR法, dqds法, mdLVs法を用いた特異値分解の方法を詳細に解説 									

○大次元疎行列の特異値分解法（木村欣司/7回 講義と演習）

- (1) C言語の基本的な文法などを解説
- (2) 大次元疎行列(大次元隣接行列)の特異値分解法を網羅的に解説
- (3) C言語を用いた Lanczos 法(G-K-L 法)の実装

〔履修要件〕

特になし

〔成績評価の方法・基準〕

「ビッグデータの可視化」、「密行列の特異値分解法」、「大次元疎行列係数の特異値分解法」で、それぞれ1つずつのレポート課題を出題し、そのレポートの内容から習熟度を判断し、さらに出席状況を加えて、総合的に評価します。

〔教科書〕

講義資料を配布。
特に定めなし。

〔参考書等〕

小山田耕二, 坂本尚久『粒子ボリュームレンダリング-理論とプログラミング』
(コロナ社)ISBN:978-4-339-02449-4
(See http://www.coronasha.co.jp/np/detail.do?goods_id=2726)

〔その他(授業外学習の指示・オフィスアワー等)〕

オフィスアワーについては担当教員の KULASIS 登録情報を参照すること。

木村欣司:kkimur@amp.i.kyoto-u.ac.jp