M3239.006800 Geometric Methods for High-Dimensional Data Analysis
Fall 2022 Syllabus

Introduction: This course introduces geometric methods for learning low-dimensional representations of high-dimensional data. After covering the fundamentals of differential geometry – specific topics include Riemannian manifolds and Lie groups, tensors, geodesics, connections and fiber bundles – we develop geometric generalizations of traditional machine learning algorithms (e.g., manifold learning, metric learning) for data that may be high-dimensional, non-Euclidean, and possess some underlying symmetry. Efficient computational algorithms to implement these methods are also introduced. We also investigate the connections between reinforcement learning and stochastic optimal control from the perspective of Itô stochastic differential equations. Geometric methods for reinforcement learning problems involving high-dimensional, non-Euclidean data are also developed.

Course Instructor: The instructor for this course is (Frank) Chongwoo Park (fcp@snu.ac.kr). The instructor will be available after each class for questions and discussion. Individual meetings can also be arranged by appointment.

Course Webpage: A course webpage will be maintained at *http://etl.snu.ac.kr*. All lecture notes, homework assignments, solutions, and announcements will be made available on the course webpage.

Prerequisites: This course is a graduate advanced topics course, and students are expected to have a working familiarity with basic machine learning and data science concepts as covered in a first-year graduate course or advanced undergraduate course. Students should also have basic background in linear algebra, optimization, and probability. Additional courses in systems and control, signal processing, and other related courses are helpful but not required. **All students should obtain prior permission from the instructor in order to register for the course**.

Course Materials: For the primary textbook we will use a draft version of the book *Manifolds, Geometry, and Machine Learning* by F.C. Park *et al* (to be published in 2023). For the material on stochastic differential equations and optimal control, we will use the *Lecture Notes on Stochastic Control* by Roger Brockett as our primary source. Other references and supplementary materials in the form of papers, monographs, and notes will be provided throughout the course.

Grading: The grading for the course will be based on in-class participation (10%), one examination (30%), 4-5 homework assignments (30%), and a course project (30%). The course project can be on any topic of your choice that makes direct use of the methods covered in the course. Additional points given for creativity and novelty of the project topic. You may work alone or in teams of two; if you work with a partner, the specific roles and contributions of each team member must be explicitly specified. A final project report must be submitted, and a 15-minute project presentation must be given. All code and data used in the project must be submitted or made available in order to verify replicability. Each student must submit a one-page project proposal by the end of October.

Lecture Topics: The lectures will be given on Mondays and Wednesdays from 17:00–18:15 in

Building 301, Room 204. **All lectures will be given in English.** Select guest lectures may also be scheduled throughout the course. For each topic, plentiful examples and case studies from machine learning and data science will be interspersed throughout to illustrate the concepts. The following list of lecture topics is tentative and subject to change:

- Motivation: coordinate systems, invariance, symmetry, and machine learning
- Curves and surfaces in $\mathbb{R}^3$
- Manifolds in $\mathbb{R}^n$
- Intrinsic Riemannian manifolds
- Tensor fields on manifolds
- Mappings between Riemannian manifolds
- Riemannian distortion and representation learning
- Lie groups: basic concepts, optimization, convolution
- Connections, gauges, and fiber bundles;
- Information geometry: statistical manifolds, geometry of normal densities
- Stochastic differential equations: basic theory, extension to manifolds and Lie groups,
- Optimal feedback control from a stochastic differential equation perspective, connections to reinforcement learning