CISC371: Nonlinear Data Analysis

Fall 2025

Detailed Description:

When we use machine-learning methods on structured numerical data, we often want to find an optimal answer. For example, if we want to model the data as being linearly related, we might want to minimize the error between the model and the data.

In this course, we will explore data where we have a single number that indicates how well the model represents the data. This number is the *objective* that we want to optimize. Some of the models that we will study include: neural networks; linear relationships within the data where we have a nonlinear objective, such as ridge regression and the "lasso"; and dividing data into two classes based on how the data is labeled, such as neural networks or a support vector machine, also known as a SVM.

A guiding principle for this course will be that linear algebra is our main way of describing and implementing nonlinear optimization of structured data. Most other courses use a "calculus-first" principle; instead, we will use a limited amount of calculus to derive the linear algebra that we need for writing code. This "algebra-first" principle will allow us to explore complicated ideas in a concise and scalable manner by using linear algebra as the "language" of structured data.

The beginning part of the course will deal with mathematical preliminaries. We will explore two simple ways to optimize a function that depends on just one variable, mainly using search methods. Next, we will explore a few of the basic ways to optimize a function that depends on a vector of variables; these methods combine the one-variable methods into ways of performing *unconstrained* optimization. The first part ends with neural networks, where we will describe and implement the back-propagation algorithm by using matrix computations instead of the usual combination of calculus terms and summation loops.

The middle part of the course will deal with *constraints*, which are requirements that we place on the acceptable solutions to our problems. We will write simple constraints as "expanded" matrix problems, and then we will convert the constraints into "contracted" matrix problems. These ideas are based on *Lagrange multipliers* and *duality*, respectively. We will briefly explore how to manage nonlinear constraints using the *KKT conditions*.

The final part of the course will deal with the application of constraints to problems that arise widely in data analysis. We will place constraints on simple regression problems that will help us to understand how ridge regression and the "lasso" solve these problems. We will explore how a single constraint – part of a family of *regularization* methods – can help us when we solve seemingly complicated problems in data modeling. The idea of duality will help us to use linear algebra to describe constrained solutions to problems of data classification. We will evaluate our results using current statistical methods for validation of regression and assessment of classification. The course ends by using kernels in an SVM to find nonlinear classifications of structured data.

Learning Outcomes:

By the end of the course, a successful student will be able to:

- **LO1:** Formulate given problems as optimization functions
- **LO2:** Synthesize data and solution methods for optimization
- **LO3:** Implement, test and evaluate optimization methods
- **LO4:** Interpret and explain methods and solutions of given problems
- **LO5:** Evaluate and critique performance of algorithms

These will be assessed by computer implementations and tests.

Textbook:

There is no required textbook for this course. We will use the instructor's notes and other sources that are available through permissions granted to the Queen's University Library.