DSCI 320: Optimization Methods in Data Science

Instructor:	Prof. Wolfgang Bangerth Weber 214 bangerth@colostate.edu
Lecture:	Monday, Wednesday, Friday, 1:00–1:50 pm Weber, room 223
Office hours:	Wednesday, 2–3 pm; or by appointment.

Catalog description Linear and non-linear programming, convex sets and functions, convex and non-convex optimization problems, duality, Newton's methods, barrier methods, linear equality and inequality constraints. Emphasis on computation methods and programming.

Motivation and background for this course Whenever words such as "best/worst/fastest/most efficient/..." appear, there is an *optimization problem*: Among all choices, we want to select one that is the most desirable in some sense. For example, when designing a plane, we might want to pick that wing shape that requires the *least* fuel to push through the air while staying aloft. In the context of data sciences, a typical application is to ask for the *best strategy* to decide whether a measurement corresponds to an object of type A or B, given that we have made many measurements of A- and B-type objects before. (Think about how you have probably become pretty good at distinguishing whether a piece of fruit is an apple or an orange, having seen many of these in your life. But how do we teach a computer how to make this classification if we give it the diameter and color of the object?)
Mathematically, all of these problems are formulated in the following way: Let's say you are giving a function $f(\mathbf{x}), \mathbf{x} \in \mathbb{R}^n$, how do you find that $\mathbf{x}$ for which $f(\mathbf{x})$ is larger or smaller than for any other $\mathbf{x}$? In other words, we are looking for an algorithm that can be implemented on a computer to find the minimum or maximum of f.
DSCI 320 is a course that introduces you to the existing algorithms and the thought processes behind them. Computer exercises will help you build intuition on how they work in practice.

Course topics The course will cover three large areas related to optimization:

1. *Theory:* In order to understand how optimization algorithms work, it is very useful to have a little bit of an understanding of the theory of optimization. This is principally because it provides us with the *intuition* to understand what kinds of situations can happen when we try to minimize a function $f(\mathbf{x})$ (e.g., a single minimum vs. many minima; why convexity is an important concept). An important part of understanding the theory is to construct *counter-examples* that teach us in what kind of situations we can't expect something to work.

2. *Nonlinear optimization:* When $f(\mathbf{x})$ is a nonlinear function of $\mathbf{x}$, then the general approach to finding its minimum (if one exists) involves approximating f by something simpler for which we can easily write down where its minimum lies; examples of this approach are the gradient descent and Newton methods. We then hope that the minimizer of this approximation is close to where f also has its minimum. We can then repeat this method and iteratively hopefully find our way to the actual minimum of f.

   The problem becomes a bit more complicated if there are *constraints* that we need to respect. We will discuss Lagrange multiplier and penalty/barrier methods for these cases.

   All theoretical considerations will be supported by homework problems that consider the practical implementation of these ideas.

3. *Linear optimization:* If, on the other hand, $f(\mathbf{x})$ is a *linear* function of $\mathbf{x}$, then an entirely separate set of methods needs to be used. We will explore the basic *simplex method* to solve such problems.

   Again, the theory will be augmented by practical exercises that implement the relevant algorithms.

Textbook I do not require you to get any particular book, and in particular will not pose homework that references a book. That said, if you want to read up on some of the material we discuss in class, the following books (in any edition you can find) cover essentially everything we do over the course of this semester:

- J. Nocedal and S. J. Wright: Numerical Optimization, Springer.
- D. Bertsimas and J. N. Tsitsiklis: Introduction to linear optimization, Athena Scientific.

Prerequisites (CS 163 or CS 164) and (MATH 151 and MATH 261) and (DSCI 369 or MATH 369). Many of the homework assignments will require you to write small programs. In general, I leave the choice of programming language to you, but if your choice is somewhat exotic or outside the realm of what a typical programmer can be expected to read, you will need to provide sufficient commentary to make the code understandable.

Webpage Homework assignments and other course information will be posted at the course webpage
`http://www.math.colostate.edu/~bangerth/teaching.html`

Exams + Grading Final grades will be determined based on the following components:

- Homework and programming assignments: 40%
- Midterm exam, date to be determined, 25%
- Either a final exam or a project (to be discussed in class), with exam date or project due date to be determined: 35%

Your minimum grade will be A, B, C, or D, for a score of 90%, 80%, 70%, and 60% over the course of the semester, respectively.
You must make arrangements in advance if you expect to miss an exam or quiz. Exam absences due to recognized University-related activities, religious holidays, verifiable illness, and family/medical emergencies will be dealt with on an individual basis. In all cases of absence from exams a written excuse is required. Ignorance of the time and place of an exam will not be accepted as an excuse for absence.

Learning Outcomes and Course Objectives Optimization algorithms are one of the most widely used tools in all of engineering to find designs that best implement desired outcomes. In the context of the data sciences, they form the basis of *classification methods* (aiming to find the *best* method to separate incoming data into one of several classes – as in the apples vs. oranges example above) and indeed of the broader *machine learning* approach (in which one wants to learn the underlying fundamentals of a set of data).
My goal for this class is to (i) provide a basic level of literacy in the theoretical background of optimization methods, as well as (ii) teach you about how they are implemented and work in practice. At the end of the semester, you will be able to identify and understand what methods to use depending on the situation; how they will likely perform; and analyze these methods in terms of properties such as speed of convergence. You will also have practice in implementing these methods on computers.

Policies *Academic integrity:* Academic integrity is integral to the success of the University and to you as a learner. Academic integrity is conceptualized as doing and taking credit for one's own work. Academic dishonesty undermines the educational experience at Colorado State University. Examples of academic dishonesty include (but are not limited to) cheating, plagiarism, and falsification. Plagiarism includes the copying of language, structure, images, ideas or thoughts of others and is related only to work submitted for credit. Cheating or any form of academic dishonesty will not be tolerated. The use of material from improperly cited or credited sources will be considered plagiarism. You are encouraged to collaborate with your classmates, unless otherwise directed, but all work intended for a grade must clearly be your work as an individual. Ignorance of the rules does not exclude any member of the CSU community from the requirements or the processes meant to ensure academic integrity.

Disabilities: Colorado State University, in compliance with state and federal laws and regulations, does not discriminate on the basis of disability in administration of its education related programs and activities. We have an institutional commitment to provide equal educational opportunities for disabled students who are otherwise qualified. Students with documented disabilities must contact The Office of Resources for Disabled Students (RDS; 970-491-6385) to make arrangements for class accommodations. It is the responsibility of the student to obtain accommodation letters from RDS and to make arrangements for the implementation of accommodations with faculty in advance. Students who believe they have been denied access to services or accommodations required by law should contact RDS (970-491-6385). Students who believe they have been subjected to discrimination on the basis of disability should contact the Office of Equal Opportunity (970-491-5836). For more information regarding disability grievance procedures, visit `http://oeo.colostate.edu`.