# Graduate Course - STOR 612 Foundations of Optimization (Fall 2020)

## **Course Overview**

This course is developed to introduce the foundations of optimization to graduate students at UNC-Chapel Hill, especially to students at the STOR department. It aims at conveying the most important theoretical elements and methods for graduate students who pursue research and career in different disciplines such as operations research, statistics, machine learning, computer science, engineering, applied math, and data science. The course covers both classical and modern topics from the basic to advanced theory and numerical methods. It starts from common mathematical models: linear programming and quadratic programming and moves to unconstrained convex minimization, nonconvex optimization, and then stochastic optimization. As usual, the background theory of convex optimization is convex sets, convex functions, and theory of polyhedra, which will be formally introduced in this course. Simplex methods, gradient descent algorithms, and Newton schemes that form the foundations of numerical optimization will be presented in this course. Common models such as LASSO, logistic regression, portfolio optimization, and empirical risk minimization will also be used as representative examples during the course. Recent advanced methods such as fast gradient, proximal-based, and variance reduced stochastic algorithms will be discussed. In addition to theory and algorithms, common optimization modeling software and computer solvers will also be introduced to students.

## Time and Place

Lectures: Tuesdays and Thursdays, 9:45AM - 11:00AM. Place: Hanes Art Center - Room 0121.

Office hours: 4:00-5:00PM on Tuesdays and Thursdays (Remotely) (Tentative).

Zoom's link: https://unc.zoom.us/j/94661800016?pwd=SlNsMGNRYU1ZaENMQmVuVmp2WUxOdz09 (Passcode: 015028).

### Instructor

Instructor: Quoc Tran-Dinh (quoctd@email.unc.edu)

Office: 333 Hanes Hall, UNC-Chapel Hill

### Course content

	Mathematical Optimization Models
Lecture 1	Mathematical optimization - Examples and basic concepts: decision
	variables, objective, constraints, feasible solutions, feasible set, optimal
	value, and optimal solutions, etc.
Lecture 2	Linear programming (LP) - Representative examples (old and new):
	production planning, blending, network flows, optimal transport, least
	absolute deviations (LAD), Wasserstein barycenter, etc.
Lecture 3	Forms of LPs and preprocessing (converting to standard form, solution
	construction, redundancy, etc)

	Mathematical Tools from Convex Analysis
Lecture 4	Convex sets, convex hulls, polyhedra, and convex cones: definitions,
<b>T</b> . <b>V</b>	examples, and basic properties
Lecture 5	Convex functions: definitions, examples, and basic properties
Lecture 6	Smooth convex functions and strongly convex functions
Lecture 7	Geometry of polyhedra and solution structures of LPs: basic feasible
	solutions, optimal solutions, and existence of solutions
	Linear Programming
Lecture 8	Mathematical aspects of simplex methods: matrix form, tableau form,
	and examples
Lecture 9	Two-phase simplex methods and complexity
Lecture 10	Applications and software for LPs: modeling software and LP solvers
Lecture 11	Dual problem of LP and its construction, and weak duality
Lecture 12	Strong duality, complementarity slackness, and Farkas' lemma
Lecture 13	Sensitivity analysis and robust LPs (if time permits)
	Quadratic Programming (QP)
Lecture 14	Mathematical formulations of convex QPs and examples (e.g., SVM,
	linear optimal control, and portfolio optimization)
Lecture 15	Simple QPs (e.g., least-squares and QPs with equality constraints) and
	constrained QPs, KKT conditions, and duality
Lecture 16	Solution methods for QPs: projected gradient, active-set methods, and
	interior-point algorithms.
	Unconstrained Optimization
	<b>Unconstrained Optimization</b> (With an emphasis on convex optimization, but also covering some non-
	<b>Unconstrained Optimization</b> ( <i>With an emphasis on convex optimization, but also covering some non-</i> <i>convex problems</i> (e.g., <i>nonnegative matrix factorization</i> ) <i>and methods</i> )
Lecture 17	<b>Unconstrained Optimization</b> ( <i>With an emphasis on convex optimization, but also covering some non-</i> <i>convex problems</i> (e.g., <i>nonnegative matrix factorization</i> ) and methods) Mathematical models and composite problems, representative examples
Lecture 17	Unconstrained Optimization (With an emphasis on convex optimization, but also covering some non- convex problems (e.g., nonnegative matrix factorization) and methods) Mathematical models and composite problems, representative examples (Empirical risk minimization, LASSO, logistic regression, nonnegative
Lecture 17	Unconstrained Optimization (With an emphasis on convex optimization, but also covering some non- convex problems (e.g., nonnegative matrix factorization) and methods) Mathematical models and composite problems, representative examples (Empirical risk minimization, LASSO, logistic regression, nonnegative matrix factorization, etc).
Lecture 17 Lecture 18	Unconstrained Optimization (With an emphasis on convex optimization, but also covering some non- convex problems (e.g., nonnegative matrix factorization) and methods) Mathematical models and composite problems, representative examples (Empirical risk minimization, LASSO, logistic regression, nonnegative matrix factorization, etc). Subdifferentials, normal cone, optimality condition, proximal operators,
Lecture 17 Lecture 18	Unconstrained Optimization (With an emphasis on convex optimization, but also covering some non- convex problems (e.g., nonnegative matrix factorization) and methods) Mathematical models and composite problems, representative examples (Empirical risk minimization, LASSO, logistic regression, nonnegative matrix factorization, etc). Subdifferentials, normal cone, optimality condition, proximal operators, and projection
Lecture 17 Lecture 18 Lecture 19	Unconstrained Optimization (With an emphasis on convex optimization, but also covering some non- convex problems (e.g., nonnegative matrix factorization) and methods) Mathematical models and composite problems, representative examples (Empirical risk minimization, LASSO, logistic regression, nonnegative matrix factorization, etc). Subdifferentials, normal cone, optimality condition, proximal operators, and projection The gradient descent method, its variants, and convergence analysis
Lecture 17 Lecture 18 Lecture 19 Lecture 20	<ul> <li>Unconstrained Optimization</li> <li>(With an emphasis on convex optimization, but also covering some non- convex problems (e.g., nonnegative matrix factorization) and methods)</li> <li>Mathematical models and composite problems, representative examples</li> <li>(Empirical risk minimization, LASSO, logistic regression, nonnegative matrix factorization, etc).</li> <li>Subdifferentials, normal cone, optimality condition, proximal operators, and projection</li> <li>The gradient descent method, its variants, and convergence analysis</li> <li>The accelerated gradient descent method and its variants</li> </ul>
Lecture 17 Lecture 18 Lecture 19 Lecture 20 Lecture 21	<ul> <li>Unconstrained Optimization</li> <li>(With an emphasis on convex optimization, but also covering some non- convex problems (e.g., nonnegative matrix factorization) and methods)</li> <li>Mathematical models and composite problems, representative examples</li> <li>(Empirical risk minimization, LASSO, logistic regression, nonnegative matrix factorization, etc).</li> <li>Subdifferentials, normal cone, optimality condition, proximal operators, and projection</li> <li>The gradient descent method, its variants, and convergence analysis</li> <li>The accelerated gradient descent method and its variants</li> <li>Proximal/projection operator calculations; proximal gradient methods;</li> </ul>
Lecture 17 Lecture 18 Lecture 19 Lecture 20 Lecture 21	<ul> <li>Unconstrained Optimization</li> <li>(With an emphasis on convex optimization, but also covering some non- convex problems (e.g., nonnegative matrix factorization) and methods)</li> <li>Mathematical models and composite problems, representative examples</li> <li>(Empirical risk minimization, LASSO, logistic regression, nonnegative matrix factorization, etc).</li> <li>Subdifferentials, normal cone, optimality condition, proximal operators, and projection</li> <li>The gradient descent method, its variants, and convergence analysis</li> <li>The accelerated gradient descent method and its variants</li> <li>Proximal/projection operator calculations; proximal gradient methods; and accelerated proximal gradient methods</li> </ul>
Lecture 17 Lecture 18 Lecture 19 Lecture 20 Lecture 21 Lecture 22	<ul> <li>Unconstrained Optimization</li> <li>(With an emphasis on convex optimization, but also covering some non- convex problems (e.g., nonnegative matrix factorization) and methods)</li> <li>Mathematical models and composite problems, representative examples</li> <li>(Empirical risk minimization, LASSO, logistic regression, nonnegative matrix factorization, etc).</li> <li>Subdifferentials, normal cone, optimality condition, proximal operators, and projection</li> <li>The gradient descent method, its variants, and convergence analysis</li> <li>The accelerated gradient descent method and its variants</li> <li>Proximal/projection operator calculations; proximal gradient methods; and accelerated proximal gradient methods</li> <li>Conjugate gradient methods for QPs and extensions (if time permits)</li> </ul>
Lecture 17 Lecture 18 Lecture 19 Lecture 20 Lecture 21 Lecture 22 Lecture 23	<ul> <li>Unconstrained Optimization</li> <li>(With an emphasis on convex optimization, but also covering some non- convex problems (e.g., nonnegative matrix factorization) and methods)</li> <li>Mathematical models and composite problems, representative examples</li> <li>(Empirical risk minimization, LASSO, logistic regression, nonnegative matrix factorization, etc).</li> <li>Subdifferentials, normal cone, optimality condition, proximal operators, and projection</li> <li>The gradient descent method, its variants, and convergence analysis</li> <li>The accelerated gradient descent method and its variants</li> <li>Proximal/projection operator calculations; proximal gradient methods; and accelerated proximal gradient methods</li> <li>Conjugate gradient methods for QPs and extensions (if time permits)</li> <li>Other advanced optimization methods: mirror descent, coordinate de-</li> </ul>
Lecture 17 Lecture 18 Lecture 19 Lecture 20 Lecture 21 Lecture 22 Lecture 23	<ul> <li>Unconstrained Optimization</li> <li>(With an emphasis on convex optimization, but also covering some non- convex problems (e.g., nonnegative matrix factorization) and methods)</li> <li>Mathematical models and composite problems, representative examples</li> <li>(Empirical risk minimization, LASSO, logistic regression, nonnegative matrix factorization, etc).</li> <li>Subdifferentials, normal cone, optimality condition, proximal operators, and projection</li> <li>The gradient descent method, its variants, and convergence analysis</li> <li>The accelerated gradient descent method and its variants</li> <li>Proximal/projection operator calculations; proximal gradient methods; and accelerated proximal gradient methods</li> <li>Conjugate gradient methods for QPs and extensions (if time permits)</li> <li>Other advanced optimization methods: mirror descent, coordinate de- scent, and conditional gradient methods (backup, if time permits)</li> </ul>
Lecture 17 Lecture 18 Lecture 19 Lecture 20 Lecture 21 Lecture 22 Lecture 23 Lecture 24	<ul> <li>Unconstrained Optimization</li> <li>(With an emphasis on convex optimization, but also covering some non- convex problems (e.g., nonnegative matrix factorization) and methods)</li> <li>Mathematical models and composite problems, representative examples</li> <li>(Empirical risk minimization, LASSO, logistic regression, nonnegative matrix factorization, etc).</li> <li>Subdifferentials, normal cone, optimality condition, proximal operators, and projection</li> <li>The gradient descent method, its variants, and convergence analysis</li> <li>The accelerated gradient descent method and its variants</li> <li>Proximal/projection operator calculations; proximal gradient methods; and accelerated proximal gradient methods</li> <li>Conjugate gradient methods for QPs and extensions (if time permits)</li> <li>Other advanced optimization methods: mirror descent, coordinate de- scent, and conditional gradient methods (backup, if time permits)</li> <li>Implementation aspects: per-iteration complexity analysis, line-search,</li> </ul>
Lecture 17 Lecture 18 Lecture 19 Lecture 20 Lecture 21 Lecture 22 Lecture 23 Lecture 24	<ul> <li>Unconstrained Optimization</li> <li>(With an emphasis on convex optimization, but also covering some non- convex problems (e.g., nonnegative matrix factorization) and methods)</li> <li>Mathematical models and composite problems, representative examples</li> <li>(Empirical risk minimization, LASSO, logistic regression, nonnegative matrix factorization, etc).</li> <li>Subdifferentials, normal cone, optimality condition, proximal operators, and projection</li> <li>The gradient descent method, its variants, and convergence analysis</li> <li>The accelerated gradient descent method and its variants</li> <li>Proximal/projection operator calculations; proximal gradient methods; and accelerated proximal gradient methods</li> <li>Conjugate gradient methods for QPs and extensions (if time permits)</li> <li>Other advanced optimization methods: mirror descent, coordinate de- scent, and conditional gradient methods (backup, if time permits)</li> <li>Implementation aspects: per-iteration complexity analysis, line-search, and specialization, etc (homework, tutorials, or recitations).</li> </ul>
Lecture 17 Lecture 18 Lecture 19 Lecture 20 Lecture 21 Lecture 22 Lecture 23 Lecture 24 Lecture 25	<ul> <li>Unconstrained Optimization</li> <li>(With an emphasis on convex optimization, but also covering some non- convex problems (e.g., nonnegative matrix factorization) and methods)</li> <li>Mathematical models and composite problems, representative examples</li> <li>(Empirical risk minimization, LASSO, logistic regression, nonnegative matrix factorization, etc).</li> <li>Subdifferentials, normal cone, optimality condition, proximal operators, and projection</li> <li>The gradient descent method, its variants, and convergence analysis</li> <li>The accelerated gradient descent method and its variants</li> <li>Proximal/projection operator calculations; proximal gradient methods; and accelerated proximal gradient methods</li> <li>Conjugate gradient methods for QPs and extensions (if time permits)</li> <li>Other advanced optimization methods: mirror descent, coordinate de- scent, and conditional gradient methods (backup, if time permits)</li> <li>Implementation aspects: per-iteration complexity analysis, line-search, and specialization, etc (homework, tutorials, or recitations).</li> <li>Newton methods, convergence analysis, and implementation</li> </ul>
Lecture 17 Lecture 18 Lecture 20 Lecture 21 Lecture 22 Lecture 23 Lecture 24 Lecture 25 Lecture 26	<ul> <li>Unconstrained Optimization</li> <li>(With an emphasis on convex optimization, but also covering some non- convex problems (e.g., nonnegative matrix factorization) and methods)</li> <li>Mathematical models and composite problems, representative examples</li> <li>(Empirical risk minimization, LASSO, logistic regression, nonnegative matrix factorization, etc).</li> <li>Subdifferentials, normal cone, optimality condition, proximal operators, and projection</li> <li>The gradient descent method, its variants, and convergence analysis</li> <li>The accelerated gradient descent method and its variants</li> <li>Proximal/projection operator calculations; proximal gradient methods; and accelerated proximal gradient methods</li> <li>Conjugate gradient methods for QPs and extensions (if time permits)</li> <li>Other advanced optimization methods: mirror descent, coordinate de- scent, and conditional gradient methods (backup, if time permits)</li> <li>Implementation aspects: per-iteration complexity analysis, line-search, and specialization, etc (homework, tutorials, or recitations).</li> <li>Newton methods, convergence analysis, and implementation</li> <li>Quasi-Newton methods and BFGS updates</li> </ul>
Lecture 17 Lecture 18 Lecture 19 Lecture 20 Lecture 21 Lecture 22 Lecture 23 Lecture 24 Lecture 25 Lecture 26	<ul> <li>Unconstrained Optimization</li> <li>(With an emphasis on convex optimization, but also covering some non- convex problems (e.g., nonnegative matrix factorization) and methods)</li> <li>Mathematical models and composite problems, representative examples</li> <li>(Empirical risk minimization, LASSO, logistic regression, nonnegative matrix factorization, etc).</li> <li>Subdifferentials, normal cone, optimality condition, proximal operators, and projection</li> <li>The gradient descent method, its variants, and convergence analysis</li> <li>The accelerated gradient descent method and its variants</li> <li>Proximal/projection operator calculations; proximal gradient methods; and accelerated proximal gradient methods</li> <li>Conjugate gradient methods for QPs and extensions (if time permits)</li> <li>Other advanced optimization methods: mirror descent, coordinate de- scent, and conditional gradient methods (backup, if time permits)</li> <li>Implementation aspects: per-iteration complexity analysis, line-search, and specialization, etc (homework, tutorials, or recitations).</li> <li>Newton methods, convergence analysis, and implementation</li> <li>Quasi-Newton methods and BFGS updates</li> <li>Stochastic optimization methods</li> </ul>
Lecture 17 Lecture 18 Lecture 19 Lecture 20 Lecture 21 Lecture 22 Lecture 23 Lecture 24 Lecture 25 Lecture 26 Lecture 27	<ul> <li>Unconstrained Optimization</li> <li>(With an emphasis on convex optimization, but also covering some non- convex problems (e.g., nonnegative matrix factorization) and methods)</li> <li>Mathematical models and composite problems, representative examples</li> <li>(Empirical risk minimization, LASSO, logistic regression, nonnegative matrix factorization, etc).</li> <li>Subdifferentials, normal cone, optimality condition, proximal operators, and projection</li> <li>The gradient descent method, its variants, and convergence analysis</li> <li>The accelerated gradient descent method and its variants</li> <li>Proximal/projection operator calculations; proximal gradient methods; and accelerated proximal gradient methods</li> <li>Conjugate gradient methods for QPs and extensions (if time permits)</li> <li>Other advanced optimization methods: mirror descent, coordinate de- scent, and conditional gradient methods (backup, if time permits)</li> <li>Implementation aspects: per-iteration complexity analysis, line-search, and specialization, etc (homework, tutorials, or recitations).</li> <li>Newton methods, convergence analysis, and implementation</li> <li>Quasi-Newton methods and BFGS updates</li> <li>Stochastic optimization methods</li> </ul>

# **Course materials**

### Lecture notes

Lecture notes and slides will be provided to students via Sakai. They must be used internally in this course. Please do not distribute these materials without the instructor's permission.

### **Reference Books**

- [B1]. R. T. Rockafellar: Convex Analysis, 1970, Princeton Univ. Press. This book is now available online for free and can be downloaded from http://www.convexoptimization.com/TOOLS/ ConvexAnalysisRockafellar.pdf.
- [B<sub>2</sub>]. D. Bertsimas and J. Tsitsiklis: Introduction to Linear Optimization, Third Edition, Athena Scientic, Belmont, Massachusetts, 1997.
- [B<sub>3</sub>]. S. Boyd and L. Vandenberghe: Convex Optimization, 2006, Cambridge Univ. Press. This book is available for free at http://stanford.edu/~boyd/cvxbook/.
- [B<sub>4</sub>]. A. Beck: First-order methods in optimization, volume 25, SIAM, 2017.
- [B<sub>5</sub>]. Y. Nesterov: Introductory lectures on Convex Optimization, 2004 or 2018. The lectures can be found at http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.693.855&rep= rep1&type=pdf.
- [B<sub>6</sub>]. D. Bertsekas, Convex Optimization Theory/Algorithms, Athena Scientific, 2009.
- [B<sub>7</sub>]. J. Nocedal and S. Wright, Numerical Optimization, Springer-Verlag, 2006.
- [B<sub>8</sub>]. G. Lan, First-order and Stochastic Optimization Methods for Machine Learning, Springer-Nature, 2020.

# Prerequisites

This course requires basic knowledge from numerical linear algebra and multivariable calculus at least from the MATH547/MATH577 level or equivalence (e.g., STOR 415). If student does not have such background, then I will highly recommend to drop the course.

# Class attendance and course website

Students are expected to attend all classes. Students are responsible for assignments or policies that are announced in class or in material handed out in class, whether or not students are in class. Students are also responsible for any material distributed electronically by email or via the course webpage. A course website is available at https://sakai.unc.edu/. Once student logs in with his/her ONYEN and password, go to the course site entitled STOR612.001.FA20. We will use it for delivering assignments, posting lecture notes and grades, and for other purposes.

# Homework

Homework will be assigned in most weeks. It will be posted at the course site on the "Assignments" page in Sakai. It will be due on dates stated in the assignment information posted in Sakai; it will be graded, and the grades returned to you, usually within one week. For Fall 2020, the homework will be submitted electronically via Sakai system, at the course site on the "Assignments" page. Students scan their homework or take pictures, make sure that the electronic version is clear and clean, and put all pages in one PDF file. Please keep the file at a reasonable size to be able to attach

into the Sakai system. Any homework not turned in at the regular class meeting due to foreseen reasons must be sent to the instructor via email before 5:00PM on the due date. Do not deposit homework in the instructor's mailbox or the instructor's office. In general, late homework will receive no credit. Occasionally, reasonable exceptions may be made, with the instructor's specific approval in each case. Verbatim copying of homework is absolutely forbidden and constitutes a violation of the Honor Code. To receive full credit on homework, students must:

- show all work neatly, write in blue or black pen or pencil, with student's name in blue or black ink on the first page;
- clearly label each problem;
- staple the entire assignment together in the correct order.

Students who believe their grade on an assignment or examination is in error can request adjustment of the grade during a period of three weeks after the due date of the item in question. The threeweek period may be shortened for the last one or two assignments of the semester. Any questions regarding homework grades should first be taken up with the grader; if these questions cannot be resolved with the grader, then feel free to discuss them with the instructor. Students are responsible for checking grade book regularly (e.g., weekly) and interacting with the instructor to prevent any mistake on grading for both homework assignments and exams during the semester.

#### Exams

There will be one in-class examination (75 minutes), scheduled in one class during the midterm week, and one final examination (three hours). The exact time of the midterm will be announced directly via Sakai. The time and place for the final exam are scheduled by the Registrar office, and can be found at https://registrar.unc.edu/academic-calendar/. This time will also be informed in class. In Fall 2020, the midterm and final exams could be changed if the university moves to remote mode. If this happens, then students will be announced in advance. Any questions regarding exam grades should be taken up with the instructor. There is no make-up exam for the midterm one. If student has an official reason such as sickness, please contact the instructor to discuss possible solutions. This arrangement should be done as early as possible before the exam if student already has a plan, or after the exam, otherwise.

### Course grade

A student's course grade will be based on the final course average, in computing which the graded work will be weighted as follows: regular homework assignments: 25%; in-class examination: 25%; final exam: 50%. No homework assignment is dropped. All questions about course registration and waitlists should be directed to Ms. Christine Keat (crikeat@email.unc.edu, Hanes 321, 919-962-2307). Again, each student is responsible for verifying his or her recorded scores (homework & in-class exams) during the semester.

The Honor Code will be observed at all times in this course. The terms of the Honor Code are set out at http://instrument.unc.edu. Please carefully check it.

#### Surveys

The instructor may launch one or two surveys to get student's feedback. These surveys are completely anonymous. Students can write their comments and suggestions to help the instructor improving his teaching progress and the quality of the course. However, students can always give direct feedback to the instructor during class to improve the quality of the course.

## Community Standards in Our Course and Mask Use

This fall semester, while we are in the midst of a global pandemic, all enrolled students are required to wear a mask covering your mouth and nose at all times in our classroom. This requirement is to protect our educational community – your classmates and me – as we learn together. If you choose not to wear a mask, or wear it improperly, I will ask you to leave immediately, and I will submit a report to the Office of Student Conduct. At that point you will be disenrolled from this course for the protection of our educational community. An exemption to the mask wearing community standard will not typically be considered to be a reasonable accommodation. Individuals with a disability or health condition that prevents them from safely wearing a face mask must seek alternative accommodations through the Accessibility Resources and Service. For additional information, see Carolina Together.