The first exam will, broadly speaking, cover all the material covered in class between Monday, February 5th and Friday, March 2nd. In this document, I go through each relevant section of the textbook and point out the things you should make sure to know from it.

To study for the exam, I suggest the following resources:

- Review the homework assignments. Moodle (https://learn.illinois.edu/) will have solution sets for each assignment each Monday. It’s worth reading over them even if you solved the problems correctly to see if you missed shortcuts or alternative approaches.

- When reviewing material from a specific section, look at the examples in the textbook, which are labeled “(x.y.z) Example.”, and try to work through or understand them on your own before checking against the book’s solution.

- The exercises at the end of each chapter are pretty good too. Some of them develop extra theory that wasn’t covered in that chapter, which occasionally means they get deeper into the material than we do in this class. If you have questions about how to solve any of the exercises, I’m happy to answer them.

2 Convex Sets and Convex Functions

2.3 Convex Functions

You should be able to show that a function is convex in all of the following ways:

- By checking that the definition of a convex function (Definition 2.3.2) applies.

- By checking that its Hessian matrix is positive semidefinite (Theorem 2.3.7).

- By building it out of simpler convex functions, using Theorem 2.3.10 to show that the result is convex. (Theorem 2.3.10 has a number of subcases, which we discussed in class.)

- By using the result that, whenever $T : \mathbb{R}^m \to \mathbb{R}^n$ is an affine transformation and $f : \mathbb{R}^n \to \mathbb{R}$ is a convex function, then $f \circ T : \mathbb{R}^m \to \mathbb{R}$ is also convex. (In particular, if $h : \mathbb{R} \to \mathbb{R}$ is convex, then functions of the form $f(x_1, \ldots, x_n) = h(x_i)$ are convex.)
2.4 Convexity and the AM-GM Inequality

You should know the AM-GM inequality (Theorem 2.4.1), paying special attention to the equality case, and how it arises from applying Jensen’s inequality to \( f(x) = -\log x \).

You should be able to apply the AM-GM inequality to solve constrained geometric minimization problems in simple cases. (For example, the problems in Exercise 15 at the end of Chapter 2.)

2.5 Unconstrained Geometric Programming

Given an unconstrained geometric program, you should be able to:

- Write down its dual program.
- Understand how the constraints in the dual program arise from the requirements of applying the AM-GM inequality to the primal objective function.
- Find the dual optimal solution in cases where it is unique.
- In cases where the dual optimal solution is not unique, at least parametrize the dual feasible set (as in Example 2.5.5(c) in the textbook) even if solving for the critical point is hard to do by hand.
- Use the dual optimal solution to solve for the primal optimal solution.

(In general, you should be sure that you understand Example 2.5.5 in the textbook very well.)

You should also know the general fact that \( g(t) \geq v(\delta) \) for all primal feasible \( t \) and dual feasible \( \delta \), both for this exam and for understanding optimization duality in broader form when we get to Chapter 5.

2.6 Convexity and Other Inequalities

We skipped this section, so you don’t have to know this material for the exam.

Affine Transformations

This is not a topic formally discussed with the textbook, and I spent a day on it only so that it would help you understand other topics in this class better. (For example, the result mentioned at the end of section 2.3 above is harder to state and to prove if you don’t know what an affine transformation is.)

There will not be questions on the exam specifically about affine transformations.
3 Iterative Methods for Unconstrained Optimization

For all of these methods, don’t worry about heavy numerical computation appearing on the final. You should be able to do the calculations in principle, assuming the numbers are not too bad.

3.1 Newton’s method

You should be able to use all versions of Newton’s method and understand their geometric interpretations (as finding roots of linear approximations, or critical points of quadratic approximations, to the original function). Be aware of cases where Newton’s method fails and what that looks like.

You should understand Theorems 3.1.4 and 3.1.5 (although the latter isn’t relevant until our discussions of descent methods).

You don’t need to worry about the discussion of how to solve systems of linear equations efficiently.

3.2 The method of steepest descent

You should be able to use the method of steepest descent. Theorems 3.2.5 and 3.2.6 are important to know; theorem 3.2.3 is kind of cool.

3.3 Beyond steepest descent

You should know the criteria 1–4 for Wolfe’s theorem (Theorem 3.3.1), as well as the theorem itself, and be able to do the following things:

1. Find a descent direction \( p^{(k)} \) of the form \(- (Hf(x^{(k)}) + \mu_k I)^{-1} \nabla f(x^{(k)})\).

2. Test which values \( t_k \) in the iterative step \( x^{(k+1)} = x^{(k)} + t_k p^{(k)} \) satisfy Criteria 3 and 4 of Wolfe’s theorem.

3.4 Broyden’s method

You should be able to use Broyden’s method (3.4.1) and understand the geometric intuition behind the update rule.