Math 484: Topics Covered in Exam 2

Misha Lavrov

October 11, 2018

The second exam will, broadly speaking, cover all the material covered in class between Monday, September 17th and Friday, October 12th, with some review of previous material. 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 solution to the homework assignments. Moodle (https://learn.illinois.edu/) will have solution sets for each assignment on the day it is returned. 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.
- By using the following result, which was mentioned in class and appears as Exercise 24 at the end of Chapter 2:

If $f(x)$ is a convex function on $\mathbb{R}$ and $\mathbf{a} \in \mathbb{R}^n, b \in \mathbb{R}$, then $f(\mathbf{a} \cdot \mathbf{x} + b)$ is a convex function on $\mathbb{R}^n$.
More generally, if $f(\mathbf{x})$ is a convex function on a convex set $C \subseteq \mathbb{R}^n$, A is an $m \times n$ matrix, and $\mathbf{b} \in \mathbb{R}^n$, then

$$f(A\mathbf{x} + \mathbf{b})$$

is a convex function on $C' = \{\mathbf{x} \in \mathbb{R}^m : A\mathbf{x} + \mathbf{b} \in C\}$.

(This is the most general form, but in practice, this result is used to verify simple statements such as "$g(x, y) = (x - y + 1)^4$ is convex because $f(t) = t^4$ is convex".)

There are a few other important ideas in this section as well:

- The notion of *strictly convex functions*.
- The first and second derivative tests for convexity: Theorem 2.3.5 and Theorem 2.3.7.
- Consequences of these definitions for optimization, such as Corollary 2.3.6.

Pay specific attention to what these results *don't* say. For example, Corollary 2.3.6 doesn't guarantee the existence of any critical points or global minimizers of a convex function f, and Theorem 2.3.7 doesn't guarantee that a strictly convex function will always have a positive definite Hessian matrix.

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

You should understand the general idea of a dual optimization problem, and how duality applies to the specific case of 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.)

2.6 Convexity and Other Inequalities

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

4 Least Squares Optimization

4.1 Least Squares Fit

You should know how to solve a least-squares optimization problem of the form

$$\min\{\|A\mathbf{x} - \mathbf{b}\| : \mathbf{x} \in \mathbb{R}^n\}$$

by solving the normal equation $A^\mathsf{T} A\mathbf{x} = A^\mathsf{T}\mathbf{b}$.

In particular, you should know how to use this procedure to find a best-fit line or, more generally, a best least-squares fit for equations of a given form.

We also talked about polynomial interpolation, which the textbook doesn't cover, but which you can review from https://en.wikipedia.org/wiki/Lagrange_polynomial. (Only the material in sections 1–4 of the article is relevant to us.)

Finally, you should know the following things about orthonormal bases:

- The definition of an orthonormal basis.
- If Q is a matrix with orthonormal columns $\mathbf{u}^{(1)}, \dots, \mathbf{u}^{(k)}$, then $Q^\mathsf{T} Q = \mathbf{I}$...
- ...and $QQ^\mathsf{T} = \mathbf{u}^{(1)}(\mathbf{u}^{(1)})^\mathsf{T} + \dots + \mathbf{u}^{(k)}(\mathbf{u}^{(k)})^\mathsf{T}$ is the projection matrix onto the span of $\mathbf{u}^{(1)}, \dots, \mathbf{u}^{(k)}$.
- How to use Gram–Schmidt process to find an orthonormal basis for the span of a given set of vectors.
- How to deduce a QR factorization of A from applying the Gram–Schmidt process to the columns of A.

4.2 Subspaces and Projections

This section deals with interpreting least-squares optimization geometrically: as the projection onto the subspace $\{A\mathbf{x} : \in\mathbb{R}^n\}$. Key facts to know about subspaces (more precisely, *vector subspaces* or *linear subspaces*)

- Their definition (Definition 4.2.1) in the textbook.

- Two ways to define a subspace V: as a set of the form $V = \{A\mathbf{x} : \in\mathbb{R}^n\}$ (which makes V the set of vectors spanned by the columns of A)...
- ...or as a set of the form $V = \{\mathbf{y} \in \mathbb{R}^n : A\mathbf{y} = \mathbf{0}\}$ (which makes V the set of vectors that satisfy the linear conditions given by the rows of A).

You should know that:

- The projection map $AA^\dagger = A(A^\mathsf{T}A)^{-1}A^\mathsf{T}$, which takes a vector $\mathbf{y}$ to its closest point in the subspace $\{A\mathbf{x} : \in\mathbb{R}^n\}$.
- This closest point $A\mathbf{x}^*$ is characterized by the property that $A\mathbf{x}^* - \mathbf{y} \perp A$ for all A of the form $A = A\mathbf{x}$.

4.3 Minimum Norm Solutions of Underdetermined Linear Systems

You should know how to solve the optimization problem

$$\min \|\mathbf{x}\| : A\mathbf{x} = \mathbf{b}$$

for the underdetermined system $A\mathbf{x} = \mathbf{b}$, and understand the key property of the solution $\mathbf{x}^*$: it satisfies $\mathbf{x}^* \perp \mathbf{y}$ for all $\mathbf{y}$ such that $A\mathbf{y} = \mathbf{0}$.

You should have an idea of what is going on geometrically when we find the minimum norm solution.

4.4 Generalized Inner Products and Norms

You should be familiar with the properties of inner products (see the lecture notes) and their characterization.

Second, you should know how to apply the generalized inner products to solve the minimum-H-norm problem.

Although it was mentioned in class, you do not need to worry about the problem

$$\min \mathbf{x}^\mathsf{T}H\mathbf{x} : A\mathbf{x} = \mathbf{b}$$

when H is not positive definite.

The book also makes a big deal out of example 4.4.5 (The Portfolio Problem) but you don't need to pay this any more attention than any other example in the textbook.

The method of Lagrange multipliers

You should be able to apply the method of Lagrange multipliers to solve problems. Aside from the lecture notes, sections 7.1 and 7.2 of the textbook are reasonable sources to look at, but they also allude to topics we haven't covered yet. Most of the exercises in Chapter 7 are useful.

Watch out for conditions on when a problem actually is guaranteed to have an optimal solution.