ECE 490: Introduction to Optimization **Fall 2018**

Lecture 26

Final Review

Lecturer: Bin Hu, Date:12/11/2018

Today we give a review for the contents covered in this course. We talked about three problems in this course:

- Unconstrained optimization:

$$\min_{x \in \mathbb{R}^p} \; f(x) \tag{26.1}$$

- Optimization with equality constraints:

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & h_i(x) = 0, \; i = 1, \dots, m \end{aligned} \tag{26.2}$$

- General constrained optimization

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & h_i(x) = 0, \; i = 1, \dots, m \\ & g_j(x) \le 0, \; j = 1, \dots, l \end{aligned} \tag{26.3}$$

We have talked about optimality conditions and also related algorithms for all the above three problems.

1. **Optimality conditions**: For(26.1), the necessary condition for a local min is simple, i.e. $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*) \ge 0$. We have also talked about the sufficient condition $\nabla^2 f(x^*) > 0$. Similarly, we talked the optimality conditions (basically the Lagrange theorem) for (26.2). See Theorem 18.1 in Lecture 18 and Section 20.3 in Lecture 20 for more discussions. For (26.3), the optimality conditions are the so-called KKT conditions. See Lecture 21 and Problem 1 in HW6 for discussions. It is important to know how to apply such optimality conditions to obtain local mins for simple optimization problems.

2. **Algorithms for unconstrained optimization**: For (26.1), we have introduced various algorithms including gradient method, Nesterov's accelerated method, Heavy-ball method, Newton's method, and BFGS method. It is important to know the update formulas for all these methods. If f is L-smooth and m-strongly convex, we know the gradient method and Nesterov's method achieve linear convergence. If f is only

convex and smooth, then the gradient method converges at a rate $O(1/k)$, which is sublinear. It is important to know the difference between sublinear convergence, linear convergence, and superlinear convergence. In addition, the stepsize rule is also important. See Problem 2 in HW2 to learn how to select momentum for Nesterov's method and Heavy-ball method. See Problem 3 in Midterm 1, Section 12.1 in Lecture 12, and Section 13.3 in Lecture 13 for more discussions on stepsize rules. Both direct line search and Armijo rule are important.

3. **Proximal operator and projection operator:** Proximal operator is an important concept for nonsmooth optimization. The convergence rate proofs for the proximal gradient method are definitely relevant. It is also important to understand how proximal operators are calculated for specific problems. For example, sometimes proximal operator just involves doing projection. See Lecture 16 and HW4 for discussions.

4. **Duality theory:** It is important to understand how to formulate dual problems for constrained optimization problems. Such dual problems may lead to important insight for the original problems. See Lecture 18 (especially Section 18.3), Lecture 22, and HW6 for discussions.

5. **Methods for constrained optimization**: We have talked about method of multipliers, ADMM, and interior point methods.See Lecture 19, HW5, and Midterm 2 for more discussions on ADMM. See Lecture 23 and HW7 for more discussions on interior point method.