Optimization Notes

Xi Tan (tan19@purdue.edu)

November 16, 2017

Contents

Chapter 1

Introduction

Quadratic optimization problems (including, e.g., least-squares) form the base of the hierarchy; they can be solved exactly by solving a set of linear equations. *Newton's method* is the next level in the hierarchy. In Newton's method, solving an unconstrained or equality constrained problem is reduced to solving a sequence of quadratic problems. *The interior-point methods*, which form the top level of the hierarchy, solve an inequality constrained problem by solving a sequence of unconstrained, or equality constrained, problems. Besides Newton's method, there are quasi-Newton, conjugate-gradient, bundle, cutting-plane algorithms, and etc.

Optimization problems can be broadly divided into two types: linear optimization and nonlinear optimization, the later of which consists of unconstrained and constrained optimization problems. Obtaining necessary and sufficient conditions is one of the central problems of nonlinear optimization. The main theory is the study of Lagrange multipliers, including the Karush-Kuhn-Tucker (KKT) theorem and its extensions. However, the theory of Lagrange multipliers is far from adequate, since it does not take into account the difficulties associated with solving the equations resulting from the necessary conditions.

Definition 1.0.1 *Global convergence analysis.* The verification that a given algorithm will in fact generate a sequence that converges to a solution point. *Local convergence analysis* or *complexity analysis.* The rate at which the generated sequence of points converges to the solution.

Chapter 2

Linear Programming

2.1 Basic Properties of Linear Programs

Definition 2.1.1 Let A be an $m \times n$ matrix and B be any *nonsingular $m \times m$* sub-matrix made up of columns of A. Then, if all $n - m$ components of x not associated with columns of B are set equal to zero, the solution to the resulting set of equations is said to be a *basic solution* with respect to the basis B. The components of x associated with columns of B are called *basic variables*.

Definition 2.1.2 If one or more of the basic variables in a basic solution has value zero, that solution is said to be a *degenerate basic solution*.

Definition 2.1.3 A feasible solution that is also basic is said to be a *basic feasible solution*; if this solution is also a degenerate basic solution, it is called a *degenerate basic feasible solution*.

Theorem 2.1.1 Fundamental Theorem of Linear Programming. Given a linear program in standard form. i) if there is a feasible solution, there is a basic feasible solution; ii) if there is an optimal feasible solution, there is an optimal basic feasible solution.

Simplex algorithm has an exponential worst-case complexity, but polynomial average-case complexity.

Chapter 3

Unconstrained Optimization

3.1 Univariate Problems (Bisection, Newton)

Bisection Method

Suppose g' is continuous on $[a_0, b_0]$ and $g'(a_0)g'(b_0) \leq 0$, then the Intermediate Value Theorem implies that there exists at least one x^* for which $g'(x^*) = 0$ and hence x^* is a local optimum of g. To find this local optimum, the Bisection Method systematically halves the interval at each iteration, by checking the product of g'.

The updating equations are

$$[a_{t+1}, b_{t+1}] = \begin{cases} [a_t, x^{(t)}], & \text{if } g'(a_t)g'(x^t) \leq 0 \\ [x^{(t)}, b_t], & \text{if } g'(a_t)g'(x^t) > 0 \end{cases} \tag{3.1}$$

and $x^{t+1} = \frac{a_{t+1} + b_{t+1}}{2}$.

Newton's Method

Suppose g twice differentiable. At iteration t, Newton's method approximates $g'(x^*)$ by the linear Taylor series expansion:

$$g'(x^*) = g'(x^{(t)}) + (x^* - x^{(t)})(g''(x^{(t)})) \tag{3.2}$$

which gives us

$$x^* = x - \frac{g'(x^{(t)})}{g''(x^{(t)})} \tag{3.3}$$

3.2 Quasi-Newton Methods

Chapter 4

Convex Optimization