Numerical Optimization

Graduate Course

Introduction

Wen Huang

School of Mathematical Sciences
Xiamen University

Compiled on February 22, 2022

Optimization Problem

Optimization problem:

$$\min_{x \in \mathcal{F}} f(x)$$

- $\mathcal{F}$ is the domain
- f is a real-valued function from $\mathcal{F}$ to $\mathbb{R}$

Recipe for Solving an Optimization Problem

Basic recipe:
Initial guess $\Longrightarrow$ If unacceptable $\Longrightarrow$ Update it until acceptable

- Optimality condition
- Algorithm

Domain and objective play important roles in optimality condition analysis and algorithm design.

Optimality conditions

- Necessary conditions: (solutions $\Rightarrow$ conditions)
- Sufficient conditions: (conditions $\Rightarrow$ solutions)
- Relate to f and $\mathcal{F}$

For example, $f \in C^2$, $\mathcal{F} = \mathbb{R}$,

- a necessary condition of x_* is $f'(x_*) = 0$
- a sufficient condition of x_* is $f'(x_*) = 0$ and $f''(x_*) > 0$

Algorithms

- Generally, we will use iterative methods

$$x_{k+1} = F(x_k, x_{k-1}, \ldots, x_{k-s})$$

- Algorithms of interest are
  - robust analytically and numerically
  - efficient in time and space complexity
  - and must converge reliably and quickly

- A locally convergent method converges to a local minimizer only when the initial point is sufficiently close to the local minimizer.
- The particular local minimizer reached is a complicated function of the initial point, the method, and the problem.
- A globally convergent method converges to a stationary point (hopefully also a local minimizer) for any initial point (or fails in one of a small number of detectable ways).
- Global optimization methods attempt (or make it very likely) to find a global minimizer. This is much more difficult in general.

Important Problem Types

- Unconstrained partly smooth/smooth continuous optimzation
- Unconstrained continuous and noisy optimization
- Nonlinear programming
- Convex programming
- Quadratic programming
- Convex Quadratic programming
- linear programming (continuous/discrete)
- linear/general integer programming
- general combinatorial programming
- Manifold optimization

This course covers

- Unconstrained smooth optimization
- Constrained smooth optimization
- Unconstrained nonsmooth optimization
- Duality
- Manifold optimization

Unconstrained smooth optimization

- Optimality conditions
- Line-search-based methods
  - For an appropriate step size
    - Step size conditions
    - Algorithms for finding a step size
    - Initial step size
  - Algorithms
    - A steepest descent method
    - Nonlinear conjugate gradient methods
    - Newton method and inexact Newton methods
    - Quasi-Newton methods
    - Limited-memory versions of quasi-Newton methods
    - Preconditioned versions of all the above methods
    - Convergence analyses
- Trust-region-based methods
  - Approximately solve a local model
  - Schemes for the local model, including Newton and quasi-Newton methods
  - Convergence analysis

Constrained smooth optimization

- Optimality conditions
  - First order optimality condition: Karush-Kuhn-Tucker (KKT) conditions
  - Second order necessary/sufficient conditions
- Linear programming
  - Standard form and geometry
  - Simplex method
  - Interior point method
- Quadratic programming
  - Active set methods
  - Interior point methods
- General constrained smooth optimization problems
  - Penalty and augmented Lagrangian methods
  - Sequential quadratic programming
  - Interior point methods

Unconstrained nonsmooth optimization

- Calculus of partly smooth functions
  - Generalization of differential
  - Calculus of the generalization of differential
- Bundle algorithms
  - Gradient sampling algorithm
  - ϵ-subgradient with quasi-Newton
- Partly smooth functions with splittable structure
  - Proximal mapping
  - Proximal gradient method
  - Invariant of proximal gradient method

Duality

- Dual problems
  - Dual problem of linear programming
  - Relation to the KKT conditions
- Primal and dual problems
  - Weak and strong duality
  - Geometric and saddle-point interpretations
- Conjugate functions
  - Definition and properties
  - Conjugate calculus
  - Subdifferentials

Manifold optimization

- Motivations
  - Blind source separation
  - Shape space
- Embedded submanifold
  - Definitions
  - Differential
  - Riemannian metric
  - Gradient
  - Retraction
- A Riemannian steepest descent algorithm
  - First order expansion
  - Line search condition
  - Convergence analysis

References I