Math 381 - Fall 2022

Jay Newby

University of Alberta

Week 13

Last Week

1. Optimization overview and theory
2. Conditioning and stability
3. Constrained optimization
4. 1D optimization methods

This Week

1. Optimization methods for higher dimensional problems (unconstrained)
    - Gradient descent and stochastic gradient descent
    - Steepest descent
    - Conjugate gradient method
    - Newton's method and quasi-Newton methods
    - Secant updating method
    - Direct Search
    - Nonlinear least squares
    - Infinite dimensional optimization problems
2. The linear assignment problem
3. The Hungarian algorithm
4. Examples of the assignment problem from image analysis (time permitting)

Infinite dimensional optimization problems

Infinite dimensional optimization problem

Let S be a suitable function space. Let $\mathcal{L} : S \to \mathbb{R}$ be a functional on S. Find

$$\min_{f \in S} \mathcal{L}[f].$$

Various constraints are often imposed on the problem.

Example:

$$\mathcal{L}[f] = \int_a^b L[x, f(x), f'(x)]dx$$

Euler Lagrange equation (1D)

$$\frac{\delta \mathcal{L}}{\delta f} = \frac{\partial L}{\partial f} - \frac{d}{dx}\frac{\partial L}{\partial f'} = 0$$

Sensitivity and conditioning

Conditioning version 1

Consider the unconstrained minimization problem. Let $\hat{x}$ be the exact minimum of f, and assume that the Hessian $H(\hat{x})$ is positive definite. Computing the minimum using only $f(x)$ is ill conditioned in the sense that if

$$|f(\hat{x} + \Delta x) - f(\hat{x})| < \epsilon,$$

then the error is

$$\|\Delta x\| = O(\sqrt{\kappa(f)\epsilon}), \quad \kappa(f) = \sup_{\|u\|=1} \frac{1}{u^T H(\hat{x}) u}.$$

Sensitivity and conditioning

Conditioning version 2

Consider the unconstrained minimization problem. Let $\hat{x}$ be the exact minimum of f and assume that the Hessian $H(\hat{x})$ is positive definite. Computing the minimum by solving $\nabla f(x) = 0$ is well conditioned in the sense that if

$$\|\nabla f(x + \Delta x)\| < \epsilon$$

then

$$\|\Delta x\| = O(\epsilon\kappa(\nabla f)), \quad \kappa(\nabla f) = \sup_{\|u\|=1} \frac{1}{\|H(\hat{x})u\|}.$$

Constrained optimality conditions

Continuous optimization problem

$$\min_{x} f(x) \quad \text{subject to} \quad g(x) = 0, \quad h(x) \leq 0.$$

where $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^m$, and $h : \mathbb{R}^n \to \mathbb{R}^p$.

Lagrange multipliers

Definition: Lagrangian

$$\mathcal{L}[x, \lambda] = f(x) + \lambda^T g(x),$$

where λ is called the Lagrange multiplier.

$$(\nabla_x, \nabla_\lambda)^T \left(f(x) + \lambda^T g(x) \right)$$

$$\begin{pmatrix} \nabla f + \lambda^T \nabla g = 0 \\ g(x) = 0 \end{pmatrix}$$

Scalar equality constraint

Let $g(x)$ be a mapping from $\mathbb{R}^n \to \mathbb{R}$ (ie $m = 1$). Define the $n \times n - 1$ matrix $Z(x)$ to have orthonormal columns such that $Z(x)\nabla g(x) = 0$. Then we need to have the condition that the matrix $Z^T BZ$ be positive definite, where the $n \times n$ matrix B has elements

$$B_{ij} = \frac{\partial^2}{\partial x_i \partial x_j} \mathcal{L}[x, \lambda]$$

Linear programming

Linear programming problem

Given a matrix $A \in \mathbb{R}^{m \times n}$ and vectors $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$.

$$\min_{x \in \mathbb{R}^n} f(x) = c^T x, \quad Ax = b, \quad x \geq 0.$$

- The feasible region is a convex polyhedron in $\mathbb{R}^n$ and the global minimum must be at one of the vertices. Brute force would require a search of $\binom{n}{m}$ terms
- Has proven useful in modeling diverse types of problems in planning, routing, scheduling, assignment, and design
- Numerous software libraries exist

Discrete optimization problems

Discrete optimization problems

$$\min_{q} f(q) \quad \text{subject to} \quad g(q) = 0, \quad h(q) \leq 0.$$

where $f : \mathbb{Z}^n \to \mathbb{R}$, $g : \mathbb{Z}^n \to \mathbb{Z}^m$, and $h : \mathbb{Z}^n \to \mathbb{Z}^p$.

The assignment problem

Description

Suppose we have n labels and n objects. Let C_{ij} be the "cost" of assigning label i to object j. Let $q_{ij} = 1$ if label i is assigned to object j and $q_{ij} = 0$ otherwise. We want

$$\min_q \sum_{ij} C_{ij} q_{ij}, \quad \sum_{j=1}^{n} q_{ij} = 1, \quad \sum_{i=1}^{n} q_{ij} = 1.$$

Some discrete problems can be solved with continuous methods

Assignment problem as a linear programming problem

$$\min_{x \in \mathbb{R}^{n \times n}} \sum_{i,j=1}^{n} C_{ij} x_{ij}$$

$$\sum_{j=1}^{n} x_{ij} = 1, \quad \sum_{i=1}^{n} x_{ij} = 1$$

$$x_{ij} \geq 0, \quad 1 - x_{ij} \geq 0$$

In general, a discrete (linear) problem can be formulated as a continuous linear programming problem if the constraint matrix has certain properties (totally unimodular)