Numerical Optimization - Summary

Andy Tràn

September 14, 2022

This will be a very personalised summary for me to use to study for the course Numerical Optimization (AIST 3010). It might be complete, it might not be, it will probably not be. For questions you can refer to andtran@ethz.ch. This summary is based of the lecture notes and should be used as a supplement to the lectures.

Contents

1 Lecture 1 - 05/09/2022: Introduction to Optimization

- **Main lecture**: Monday 12:30 - 2:15, Wednesday 5:30 -6:15 (only ESTR3112)
- **Tutorial lecture**: Wednesday 4:30-5:15
- **Prereq**: Multivariable calculus, linear algebra
- **Course materials:** Homework and exam questions solely based on lecture notes.
- **Email**: farnia@cse.cuhk.edu.hk
- **Office Hour**: Tuesday 2-3 pm, SHB Building, Office 918
- **Learning goals:**
  1. Formulating optimization problems belonging to standard optimization categories for engineering and AI tasks
  2. Applying standard optimization algorithms to solve linear and convex programming tasks
  3. Implementing standard optimization algorithms over Python
- **Grading**: Homework 0.20, midterm 0.30, final .50, participation additionally 0.05

2 Tutorial 1 - 07/09/2022: Introduction to Optimization

Example 1 (Transportation problem) *we want to minimize the total cost of transporting a commodity from m factories to n stores. We have to following constraints:*

- *factory i can supply at most a_i units of the commodity*
- *store j needs at least b_j units of commodity*
- *the cost of shipping from factory i to store j is $c_{i,j}$ per unit*

We get the following system

- *Optimization variables: $x_{i,j}$, the amount of units from fac i to store j*
- *Objective function: $\sum_{i,j} c_{i,j} x_{i,j}$*
- *Constraint function: $\sum_i x_{i,j} \leq a_i$, $\sum_j x_{i,j} \geq b_j$,, $x_{i,j} \geq 0$*

For above problem, there is no analytical solution, only an interative solution.

Example 2 (Manufacturing task) *we want to maximize the profit of producing n products from m raw materials, given that*

- *We have a profit of c_i per unit of Product i*
- *We have b_j available units of raw material j*
- *We need $a_{i,j}$ units of raw material j for manufacturing one unit of i*

We get the following system

- *Optimization variable:* x_i*, amount of units per product* i
- *Objective function:* $\sum_i c_i x_i$
- *Constraint function:* $\sum_i x_i a_{i,j} \leq b_j$ *for all* j

Obviously you have to define what the allowed values are for i and j, which is left as an exercise to the reader.

Example 3 (Sorting task) *given real numbers* $c_1, ..., c_n \in \mathbb{R}$ *we want to find the* k *smallest numbers*

- *Optimization variable: For every* $1 \leq i \leq n, x_i = \begin{cases} 1 & \text{if } c_i \text{ is among the smallest } k \\ 0 & \text{otherwise} \end{cases}$
- *Objective function:* $\sum_i^n = x_i c_i$
- *Constraint function:* $\sum_i^n x_i = k$, $x_i(1 - x_i) = 0$ *for all* i

3 Tutorial 2 - 14/09/2022: Vectors

3.1 Vectors

A vector $x = [x_1, ..., x_n]$ is a collection of numbers, arranged in a column or a row, which can be thought of as the coordinates of a point in a n-dimensional space.

- Addition: is defined elementwise, given the dimensions are the same
- Scalar product: element wise multiplication with a scalar.

Note: we by default assume a vector follows a column-representation, with real values. For row format we can just transpose.

Definition 4 (Linearly independent) *Given we have n vectors, we call them l.i. when* $c_1 x_1 + ... + c_n x_n = 0$ *only has one solution, namely all* $c_1, ..., c_n = 0$

Definition 5 (Basis) *For a subspace* $S \in \mathbb{R}^d$*, is a set of l.i. vectors* $B = [x_1, ..., x_m]$ *such that every vector* $x \in S$ *is a linear combination of the vectors in B.*
Standardbasis: defined where $0 \leq i \leq m$*, where i'th coordinate of* $e_i \in S_b$ *is 1 and all the others 0.*

Definition 6 (Euclidean length) $x := \sqrt{x_1^2 + ... + x_n^2}$

Definition 7 (Norm function) *Properties of a norm function*

1. $||x|| \geq 0$*, equal to 0 only when* $x = 0$
2. *For every* $c \in \mathbb{R}, ||cx|| = |c|||x||$
3. *For every* $x, y \in \mathbb{R}^d, ||x + y|| \leq ||x|| + ||y||$

Definition 8 (l_p-norm) *For every* $p \geq 1$ *we define* l_p *norm* $|| \cdot ||_p : \mathbb{R}^d \to \mathbb{R}$ *as* $||x||_p = (|x_1|^p + ...|x_n|^p)^{\frac{1}{p}}$

Note: $l_\infty = \max\limits_{1 \leq i \leq d} x_i$
Theorem: $l_p - norms$ are decreasing in $p \geq 1$, so

$$1 \leq p \leq q \leq \infty \Rightarrow ||x||_p \geq ||x||_q.$$