

SUGGESTED EXERCISES: MODERN OPTIMIZATION

JEAN-LUC BOUCHOT

1. INTRODUCTION

2. CALCULUS-FREE OPTIMIZATION

Homework 1. Prove the following proposition:

Proposition 2.1 (Homework). *The convexity of a function can equivalently be written as: If f is convex, then for any $\alpha_1, \dots, \alpha_n$ positive numbers such that $\alpha_1 + \dots + \alpha_n = 1$, it holds*

$$f\left(\sum_i \alpha_i x_i\right) \leq \sum_i \alpha_i f(x_i).$$

Homework 2. Prove the following result:

Proposition 2.2 (Generalized AGM Inequality – Homework). *Let x_1, \dots, x_n be positive numbers and let $\alpha_1, \dots, \alpha_n$ be positive numbers such that $\alpha_1 + \dots + \alpha_n = 1$. Then the GAGM inequality reads*

$$\prod_{i=1}^n x_i^{\alpha_i} \leq \sum_{i=1}^n \alpha_i x_i.$$

Homework 3. Use the generalized AGM to prove Hölder's inequality:

Proposition 2.3 (Hölder's inequality – Homework). *For two sequence of numbers $\{a_k, 1 \leq k \leq n\}$ and $\{b_k, 1 \leq k \leq n\}$, Hölder's inequality reads, for p, q such that $\frac{1}{p} + \frac{1}{q} = 1$ (we say that p and q are Hölder conjugates)*

$$\sum_{k=1}^n |c_k d_k| \leq \|c\|_p \|d\|_q$$

with equality if and only if

$$\left(\frac{|c_k|}{\|c\|_p}\right)^p = \left(\frac{|d_k|}{\|d\|_q}\right)^q, \quad \text{for all } k.$$

And use this result to prove the triangle inequality for general p norms (so called Minkowski's inequality).

3. GEOMETRIC PROGRAMMING

Homework 4. Minimize the function $f_0(x, y) = \frac{1}{xy} + xy + x + y$ for $x > 0$ and $y > 0$.
(You might need some help of a computer.)

4. CONVEX FUNCTIONS AND ANALYSIS– REVIEW

Homework 5. Let $f : \mathbb{R}^d \rightarrow \mathbb{R}$ be a convex function defined on the whole of \mathbb{R}^d . Show that if f is bounded above, then f is constant.

Homework 6. Show that if $\text{dom}(f)$ is closed, then it is not necessarily continuous.

Homework 7. Show that the ℓ_1 norm defined as $\|x\|_1 = \sum_{i=1}^d |x_i|$ is a convex function.

Homework 8. Show that the following function is convex:

$$f : \begin{cases} \mathbb{R}^n & \rightarrow \mathbb{R} \\ x = (x_1, \dots, x_n) & \mapsto f(x) = \log\left(\sum_{i=1}^n e^{x_i}\right). \end{cases}$$

Homework 9. Let f be defined as

$$f(x, y) = x^2(1 - y^2) + y^2(1 - x^2)$$

for $(x, y) \in [-1, 1]^2$.

- (1) Show that the function f is convex along the canonical x axis. (Make sure to look at all directions parallel to the x axis, not just the one going through the point $(0, 0)$)
- (2) Show that the function f is convex along the canonical y axis. (same remark)
- (3) Show that the function f is not convex.

This example shows that canonical along all directions is not sufficient to show convexity globally.

Homework 10. Let $S_{\geq 0}^n := \{A \in \mathbb{R}^{n \times n} : A^T = A \text{ and } A \succcurlyeq 0\}$ be the set of symmetric positive semidefinite matrices. Show that $S_{\geq 0}^n$ is a convex cone. For the case $n = 2$, characterize its boundary as a surface in dimension 3.

Homework 11. Let $K := \{\alpha \in \mathbb{R}^{n+1} : \sum_{i=0}^n \alpha_i x^i \geq 0, \forall x \in [0, 1]\}$ be the set of (coefficients of) nonnegative polynomials on $[0, 1]$. Show that K is a convex cone.

Homework 12. Let C and D be two disjoint subset of \mathbb{R}^d . Consider the set of separating hyperplanes $(a, b) \in \mathbb{R}^{d+1}$ such that $a^T x - b \leq 0$ for all $x \in C$ and $a^T x - b \geq 0$ for all $x \in D$. Show that this set is convex.

Homework 13. Compute the Fenchel conjugate functions of

- (1) Negative entropy: $f(x) = x \ln(x)$.
- (2) ℓ_1 norm: $f(x) = \|x\|_1$.
- (3) $f(x) = x^p$ for some $p > 1$.
- (4) $f(x) = \max_{1 \leq i \leq d} |x_i|$.

Homework 14. Show that a geometric programming problem can be expressed as a convex problem in standard form.

Homework 15. Express the first order condition when the inequality constraints are just positivity of the coordinates.

5. GRADIENT DESCENT ALGORITHMS

Homework 16. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be the function defined as

$$f(x) = \frac{2}{3}|x|^{3/2}.$$

Define a step size $\gamma > 0$.

- (1) Show that f is convex.
- (2) Define $x^* = \left(\frac{2}{\gamma}\right)^2$. Explain the sequence of iterates given by the vanilla gradient descent starting with $x^0 = x^*$.
- (3) Consider now $x^0 \in (0, x^*)$. What can be said about the sequence of iterates?
- (4) Conclude with regards to the use of Vanilla gradient descent without care.