CS 726 - Fall 2020

Homework #1

Due : 09/21/2020, 5pm

Zijie Zhang

September 20, 2020

Question 1

Proof. When $\boldsymbol{x} = (1, 0, 0, \cdots, 0)$,

$$||x||_q = \left(\sum_{i=1}^{n} |x_i|^q\right)^{\frac{1}{q}} = 1$$

$$||x||_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} = 1$$

$$||x||_q = ||x||_p$$

When $\boldsymbol{x} = (1, 1, 1, \cdots, 1)$,

$$||x||_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} = d^{\frac{1}{p}}$$

$$d^{\frac{1}{p}-\frac{1}{q}}||x||_q = d^{\frac{1}{p}-\frac{1}{q}} \left(\sum_{i=1}^{n} |x_i|^q\right)^{\frac{1}{q}} = d^{\frac{1}{p}}$$

$$||x||_p = d^{\frac{1}{p}-\frac{1}{q}}||x||_q$$

□

Question 2

Proof. Assume that $p > 2 > q$

$$||x||_p \geqslant ||x||_2 \geqslant d^{\frac{1}{2}-\frac{1}{p}}||x||_p$$

$$||x||_2 \geqslant ||x||_q \geqslant d^{\frac{1}{q}-\frac{1}{2}}||x||_2$$

$$L_2 \geqslant \frac{||\nabla f(x) - \nabla f(y)||_2}{||x-y||_2} \geqslant \frac{d^{\frac{1}{2}-\frac{1}{q}}||\nabla f(x) - \nabla f(y)||_q}{d^{\frac{1}{2}-\frac{1}{p}}||x-y||_p} \geqslant L_p \cdot d^{\frac{2}{p}-1}$$

If $p = 1$, $L_2 \geqslant d \cdot L_p$. □

Question 3

Proof. If α_1 and α_2 are two aribitray nonnegative real numbers such that $\alpha_1 + \alpha_2 = 1$ then the convexity of f implies

$$\forall x_1, x_2,\ f(\alpha_1 x_1 + \alpha_2 x_2) \leqslant \alpha_1 f(x_1) + \alpha_2 f(x_2)$$

This can be easily generalized: if $\alpha_1, \cdots, \alpha_n$ are nonnegative real numbers such that $\alpha_1 + \cdots + \alpha_n = 1$, then

$$f(\alpha_1 x_1 + \cdots + \alpha_n x_n) \leqslant \alpha_1 f(x_1) + \cdots \alpha_n f(x_n)$$

for any $x_1, \cdots, x_n$.
By induction:
The statement is true for $n = 2$. Suppose it is also for some n, one needs to prove it for $n+1$. At least one of the α_i is strictly positive, say α_1: therefore by convexity inequality:

$$\begin{aligned} f\left(\sum_{i=1}^{n+1} \alpha_i x_i\right) &= f\left(\alpha_1 x_1 + (1-\alpha_1)\sum_{i=2}^{n+1} \frac{\alpha_i}{1-\alpha_1} x_i\right) \\ &\leqslant \alpha_1 f(x_1) + (1-\alpha_1) f\left(\alpha_1 x_1 + (1-\alpha_1)\sum_{i=2}^{n+1} \frac{\alpha_i}{1-\alpha_1} x_i\right) \end{aligned}$$

Since

$$\sum_{i=2}^{n+1} \frac{\alpha_i}{1-\alpha_1} = 1$$

one can apply the induction hypotheses to the last term in the previous formula to obtain the result, namely the finite form of the Jensen's inequality. The finite form can be rewritten as:

$$f\left(\sum_{i=1}^{k} \alpha_i x_i\right) \leqslant \sum_{i=1}^{k} \alpha_i f(x_i)$$

□

Question 4

Proof.

(i) By the lower semicontinuity, If f is lower semi-continuous at y, i.e.

$$\liminf_{\alpha \to 0} f(\alpha x + (1-\alpha)y) \geqslant f(y)$$

By convexity,

$$f(\alpha x + (1-\alpha)y) \leqslant \alpha f(x) + (1-\alpha) f(y)$$

We have,

$$f(y) \leqslant \liminf_{\alpha \to 0} \left(\alpha f(x) + (1-\alpha) f(y)\right)$$
$$f(y) \leqslant f(x) = -\infty$$

So, f is not real-valued anywhere.

(ii) Suppose on the contrary that f is not constant, that is, there exists $x_0, h \in \mathbb{R}^d$ such that $f(x_0) < f(x_0+h)$. Let $g : \mathbb{R}^d \to \bar{\mathbb{R}}$ by

$$g(t) := f(x + th)$$

for $t \in \mathbb{R}$. Note that $\mathrm{dom}(g) = \mathbb{R}$.

$$l(t) = g(0) + t[g(1) - g(0)] = f(x_0) + t[f(x_0 + h) - f(x_0)]$$

for any $t \geqslant 1$,

$$f(x_0+h) = f\left(\frac{1}{t}(x_0+th) + \left(1-\frac{1}{t}\right)x_0\right) \leqslant \frac{1}{t}f(x_0+th) + \left(1-\frac{1}{t}\right)f(x_0)$$

$$t(f(x_0+h)-f(x_0)) \leqslant f(x_0+th) - f(x_0)$$

so $l(t) \leqslant g(t)$ for $t \geqslant 1$. Since $l(\cdot)$ is an affine function with positive slope $f(x_0+th)-f(x_0)$, it is unbounded above.
Consequently,$g(\cdot)$ cannot be bounded above either, contradicting the given assumptions on g.
Therefore, f must be constant.

□

Question 5

Proof. This can be verified by using the definition of convex function and convex set.

- $\Rightarrow$ Suppose $(x, a_1),\ (y, a_2) \in epi(f)$ then $f(x) \leqslant a_1,\ f(y) \leqslant a_2$
  For any $\lambda \in [0,1]$, by convexity of f,

  $$f(\lambda x + (1-\lambda)y) \leqslant \lambda f(x) + (1-\lambda)f(y) \leqslant \lambda a_1 + (1-\lambda)a_2$$

  This implies that

  $$\lambda(x, a_1) + (1-\lambda)(y, a_2) \in epi(f)$$

  Hence $epi(f)$ is convex.

- $\Leftarrow$ Let $x,\ y \in \mathbb{R}^d$, since $(x, f(x))$ and $(y, f(y))$ lie in $epi(f)$ by convexity of epigraph set, we have $\forall\lambda \in [0,1]$

  $$(\lambda x + (1-\lambda)y,\ \lambda f(x) + (1-\lambda)f(y)) \in epi(f)$$

  By definition, $f(\lambda x + (1-\lambda)y) \leqslant \lambda f(x) + (1-\lambda)f(y))$.
  Hence, function f is convex.

□

Question 6

Proof. By **Theorem 2.1**, we know

$$f(y) = f(x) + \int_0^1 \langle \nabla f\left(x + t(y-x)\right), y - x\rangle dt$$

thus, $f(y) - f(x) - \langle \nabla f(x), y-x\rangle = \int_0^1 \langle \nabla f\left(x+t(y-x)\right) - \nabla f(x), y-x\rangle dt$
By the convexity of f, the integral is nonnegative. So, we proved

$$f(y) \geqslant f(x) + \langle f(x), y-x\rangle$$

□

Question 7

Proof. A is symmetric therefore diagonalisable in an orthonormal basis of $\mathbb{R}^d$. Let (e_i) be that basis and write

$$x = \sum_i x_i \cdot e_i,\ Ae_i = \lambda_i e_i$$

$$f(x) = \frac{\sum_i \lambda_i x_i^2}{\sum_i x_i^2}$$

Then,

$$\begin{aligned} x^T A x &= x^T A \sum_i x_i e_i \\ &= x^T \sum_i x_i \lambda_i e_i \\ &= \sum_i \lambda_i x_i^2 \end{aligned}$$

Thus,

$$\lambda_1 \leqslant f(x) \leqslant \lambda_d$$

So we proved,

(i)

$$x^T A x \geqslant \lambda_1 ||x||_2$$

(ii)

$$x^T A x \leqslant \lambda_d ||x||_2$$

□

Question 8

Proof. A is positive semidefinite means $x^T A x \geqslant 0$ for all $x \in \mathbb{R}^d$.
In **Question 7** we have prove the symmetric matrix A satisfies $x^T A x \geqslant \lambda_1 ||x||_2$ for all $x \in \mathbb{R}^d$.
So, we just need the smallest eigenvalue of matrix A is nonnegative.
By **LDL** decomposition, we know

$$L^{-1} = \begin{bmatrix} 1 & 0 & & & & & \\ 1/2 & 1 & 0 & & & & \\ 1/3 & 2/3 & 1 & 0 & & & \\ \vdots & & & & \ddots & & \\ 1/d-1 & 2/d-1 & 3/d-1 & \cdots & d-2/d-1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

Then

$$D = L^{-1} A {L^{-1}}^T$$

Obviously, the eigenvalues of A are all nonnegative.
By the way, $\det(A) = 0$, 0 is the smallest eigenvalue of matrix A. □