Assignment 1

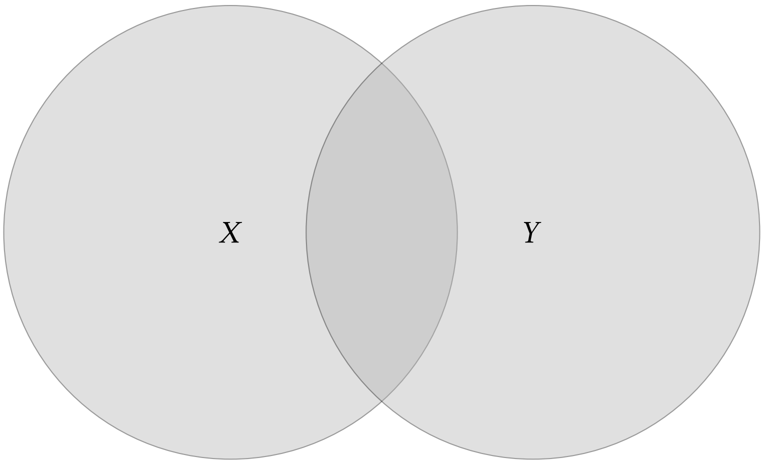
Dmitrii, Maksimov
maksimov.dmitrii.m@gmail.com

February 18, 2023

Exercise 1

Let $X, Y \subseteq \mathbb{R}^n$ and convex, then:

a) $X \cup Y$ is not consistently convex
In the set below $\exists_z : z = \alpha x + (1-\alpha)y \notin X \cup Y$, where $x, y \in X \cup Y, \alpha \in [0, 1]$



b) $X \times Y = \{(x, y) | x \in X, y \in Y\}$ is convex
Let $(x_1, y_1), (x_2, y_2) \in X \times Y$ and $z = \alpha(x_1, y_1) + (1-\alpha)(x_2, y_2)$, then
$z = (\alpha x_1 + \alpha(1-\alpha)x_2, \alpha y_1 + \alpha(1-\alpha)y_2)$. Since X, Y are convex, then $z \in X \times Y$

c) $\alpha X + \beta Y = \{\alpha x + \beta y) | x \in X, y \in Y, \alpha, \beta \in \mathbb{R}\}$ is convex
αX and βY are affine transformations and $\alpha X + \beta Y$ is Minkowski sum. Since all of these transformations preserve convexity the result set is also convex.

d) $\alpha X = \{\alpha x | x \in X, \alpha \in \mathbb{R}\}$ is convex, affine transformation

e) $X^c = \{x \in \mathbb{R}^n | x \notin X\}$ is not convex. You can imagine the complement of the X from the plot above.

Answer: a, e

Exercise 2.1

Let $f(x) = \log(x^T Ax)$, then
$\nabla f(x) = \dfrac{(A^T + A)x}{x^T Ax}$. **Question: why can't we write like this:** $df(x) = \dfrac{(A^T + A)xdx}{x^T Ax}$**? The same question for all differentials below.**

Answer: b

Exercise 2.2

Let $f(x) = \frac{1}{p}||x||_2^p$, then:

- $\nabla f(x) = \frac{d(\frac{1}{p}(x^T x)^{\frac{p}{2}})}{dx} = ||x||_2^{p-2}x$, the gradient dimension of a scalar function is n - correct
- $\nabla^2 f(x) = \frac{d\nabla f(x)}{dx} = (p-2)||x||_2^{p-4}xx^T + ||x||_2^{p-2}I$, the hessian dimension of a scalar function is $n \times n$ - correct

a) False

b) True

c) False since hessian is a scalar here

d) False

e) True

f) False

Answer: b, e

Exercise 2.3

Let $f(x) = \frac{1}{n}\sum_{i=1}^n \log(1 + \exp(a_i^T x)) + \frac{\mu}{2}||x||_2^2$, then:

- $\nabla f(x) = \frac{1}{n}\sum_{i=1}^n \frac{\exp(a_i^T x)}{1+\exp(a_i^T x)}a_i + \mu x$
- $\nabla^2 f(x) = \frac{1}{n}\sum_{i=1}^n \frac{\exp(a_i^T x)(1+\exp(a_i^T x)a_i a_i^T - (\exp(a_i^T x))^2 a_i a_i^T}{(1+\exp(a_i^T x))^2} + \mu I = \frac{1}{n}\sum_{i=1}^n \frac{\exp(a_i^T x)}{(1+\exp(a_i^T x))^2}a_i a_i^T + \mu I$

Answer: no, missed square in the first term

Exercise 3

a) $f(x) = (\sum_{i=1}^n e^{x_i})$. Since $\frac{d^n(e^x)}{dx^n} = e^x$ and $e^x > 0 \Rightarrow \nabla^2 f(x) \succeq 0$

b) $f(x) = \frac{||Ax-b||^2}{1-x^T x}$. Since $x^T x = ||x||^2 < 1$, then $f(x) \geq ||Ax - b||^2 . \nabla||Ax - b||^2 = 2A^T A \succeq 0$, since $A^T A$ PSD matrix. the proof looks crazy :)

c) don't know

d) don't know if $g(x,t) = t^2 - x^T x$ would be concave, we can prove it. This is because $-\log$ - convex, non-increasing function, but $g(x,t)$ - convex

e) From 2.3 $\nabla^2 f(x) = \frac{1}{n}\sum_{i=1}^{n} w_i \frac{\exp(a_i^T x)}{(1+\exp(a_i^T x))^2} a_i a_i^T + \mu I$. Since eigenvalues of $a_i a_i^T \geq 0$ and other constants ≥ 0 $\nabla^2 f(x) \succeq 0$.

f) don't know

Exercise 4

$$d(f(x)) = \begin{cases} [-c, c], & c^T x = 0, \\ sign(c^T x)c, & otherwise. \end{cases}$$

Exercise 5

$$d(f(x)) = \begin{cases} [-1, 1], & x = 0, \\ sign(x), & otherwise. \end{cases}$$

Exercise 6

$f(x) = ||x||_2^2 \Rightarrow \mu I \preceq \nabla^2 f(x) \preceq LI \Rightarrow L = 2, \mu = 2$

Exercise 3 addition

$$d(f(x)) = \begin{cases} (-\infty, cos(x_0)] & x = 0, \\ \varnothing & x \in (0, x_0) \\ cos(x) & x \in [x_0, 3\pi/2) \\ [0, +\infty) & x = 3\pi/2 \end{cases}$$