Exercise 1

Compute the gradient of the following functions $L\colon \mathbb{R}^3 \to \mathbb{R}$.

1. $L(w) = \|w\|_2^2$
2. $L(w) = w_1 + w_2 + w_3$
3. $L(w) = w_1 \cdot w_2 \cdot w_3$

1. $L(w) = \|w\|_2^2$
$= w^T w = w\,w$
$\nabla L(w) = w + w$
$= 2w$

2. $L(w) = w_1 + w_2 + w_3$
$\nabla L(w) = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$

3. $L(w) = w_1 \cdot (w_2 \cdot w_3)$
$\nabla L(w) = \begin{pmatrix} w_2 w_3 \\ w_1 w_3 \\ w_1 w_2 \end{pmatrix}$

Exercise 2

Which of the following sets are convex?

1. $\mathbb{R}$
2. $\mathbb{Q}$ – all convex
3. $\mathbb{Z}$
4. $\{w \in \mathbb{R}^d \mid \|w\|_2 \le 1\}$
5. $\{w \in \mathbb{R}^d \mid \|w\|_2 = 1\}$ → surface of unit ball → not convex since line betwen points along the surface would leave the set definition
6. $\{w \in \mathbb{R}^d \mid \|w\|_1 \le 1\}$ → closed unit n-ball
↳ all lines from points within the n-ball are within the set

Exercise 3

Which of the following functions $L\colon \mathbb{R}^3 \to \mathbb{R}$ are convex?

1. $L(w) = \|w\|_2^2$
2. $L(w) = w_1 + w_2 + w_3$
3. $L(w) = w_1 \cdot w_2 \cdot w_3$ → p.s.d. : convex
4. $L(w) = w^\top \operatorname{diag}(c) w$, with c a positive vector and $\operatorname{diag}(c)$ being the matrix with all zeros but c on its diagonal.

convex: $\nabla_2 L(w) \geq 0$

1. $L(w) = \|w\|_2^2$
$\nabla L(w) = 2w$
$\nabla_2 L(w) = 2 > 0$
↳ convex

2. $L(w) = w_1 + w_2 + w_3$
$\nabla L(w) = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$
$\nabla_2 L(w) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ → convex

3. $L(w) = w_1 \cdot w_2 \cdot w_3$
$\nabla L(w) = \begin{pmatrix} w_2 w_3 \\ w_1 w_3 \\ w_1 w_2 \end{pmatrix}$
$\nabla_2 L(w) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ → convex

Exercise 4

Prove the following theorem from the lecture.

Theorem. *Every local minimum of a convex function is also a global minimum.*

1. Assume x^* as a local minimum:
$f(x^*) \le f(x)$ $\forall x \in N$, where N is a neighborhood around x

2. Assume x^* is not a global minimum:
$\exists x'$, where $f(x') < f(x^*)$ for $x' \in dom(f)$

see

3. Proof contradiction:
f is convex so the following (should) hold:
$f(\lambda x^* + (1-\lambda)x') \le \lambda f(x^*) + (1-\lambda) f(x')$
i.e. → the line segment connecting x^* and x' lies above f.

3.1. local minimum property:
$f(x^*) \le f(\underbrace{\lambda x^* + (1-\lambda)x'})$
where λ is small enough so that this term is within neighborhood N (ie. $\in N$)

3.2. contradiction:
$f(x^*) \le \lambda f(x^*) + (1-\lambda) f(x') < \lambda f(x^*) + (1-\lambda) f(x^*) = f(x^*)$

Contradiction shows that a local minimum x^* is also a global one.