Tutorial 4 : Gradient descent and theoretical properties

Exercise 1 (Unproved proposition)**.** Given a C^2 function f. Prove that the following statements are equivalent:

1. f is L-smooth.
2. For any $x \in \mathbb{R}^d$, $\|\nabla^2 f\|_{\text{op}} \leq L$ where $\| \cdot \|_{\text{op}}$ of a symmetric matrix is defined by:

$$\|\mathbf{A}\| = \max_{i=1,\ldots,n} |\lambda_i(\mathbf{A})| \qquad (\lambda_i \text{ is the } i\text{th eigenvalue of } \mathbf{A}).$$

3. $-L\mathbf{I} \preceq \nabla^2 f(x) \preceq L\mathbf{I}, \forall x \in \mathbb{R}^n$.

Exercise 2 (Chebyshev polynomials)**.** For a given $n \in \mathbb{N}$, the Chebyshev polynomial of the first kind is given as the polynomial T_n such that:

$$T_n(\cos\theta) = \cos(n\theta).$$

Answer the following questions:

1. Prove that T_n can be defined alternatively as:

$$\begin{aligned} T_0(x) &= 1, \\ T_1(x) &= x, \\ T_{n+1}(x) &= 2xT_n(x) - T_{n-1}(x). \end{aligned}$$

2. Consider $f(x) = \frac{1}{2^{n-1}} T_n(x)$. Prove that the leading coefficient of $f(x)$ is 1.
3. Prove that $\max_{x \in [-1,1]} |f(x)| = \frac{1}{2^{n-1}}$.
4. Prove that for any polynomial P of degree n whose leading term is 1, $\max_{x \in [-1,1]} |f(x)| \geq \frac{1}{2^{n-1}}$. Hint: Assume a "bad" polynomial ω exists: Consider the function $g = f - \omega$. What is the degree of g? And what is the lower-bound of the number of roots of g?

Exercise 3 (Worst quadratic function)**.** The function that we used to prove the lower-bound of first-order methods is actually quadratic. To understand more about this function, consider a slightly more simple function:

$$f_k(x) = \frac{1}{2}\left[x_1^2 + \sum_{i=1}^{k-1}(x_i - x_{i+1})^2 + x_k^2\right] = \frac{1}{2}x^\top \mathbf{A}_k x.$$

Answer the following question:

1. Prove that $\mathbf{A}_k$ takes the following form:

$$\mathbf{A}_k = \left(\begin{array}{c|c} \begin{pmatrix} 2 & -1 & 0 & & & \\ -1 & 2 & -1 & & 0 & \\ 0 & -1 & 2 & & & \\ & \cdots & & & \cdots & \\ & 0 & & -1 & 2 & -1 \\ & & & 0 & -1 & 2 \end{pmatrix} & \mathbf{0}_{k\times(n-k)} \\ \hline \mathbf{0}_{(n-k)\times k} & \mathbf{0}_{(n-k)\times(n-k)} \end{array}\right)$$

2. Prove that $\mathbf{A}_k$ has $(n-k)$ zero eigenvalues and k positive eigenvalues given by:

$$2 - 2\cos\left(\frac{\pi j}{k+1}\right) = 4\sin^2\left(\frac{j\pi}{2(k+1)}\right), \quad j = 1, \dots, k$$

   Hint: You might want to use the fact that the sequence $x_0 = 1, x_1 = 2\cos(\theta), x_{k+1} = 2\cos(\theta)x_k - x_{k-1}$, then $x_k = \frac{\sin((k+1)\theta)}{\sin(\theta)}$.

3. Is the function μ-strongly convex ($\mu > 0$)? If yes, for which value of μ.

4. Is the function L-smooth? If yes, for which value of L.

5. What are the minimizer of the function $f_k, k = 1, \dots, n$?

6. For a function f that is μ-strongly convex and L-smooth, remind that we have linear convergence of f, i.e.,

$$f(x_k) - f(x^\star) \leq O\left(\left(1 - \frac{\mu}{L}\right)^k\right)$$

   Does this contradict the lower-bound that we establish in the lecture?