Exercice 1 (GD/HB for quadratic function). Let $x_\star$ be any vector, H a positive semi-definite symmetric matrix and f the convex quadratic function defined as

$$f(x) = \frac{1}{2}(x - x_\star)^T H(x - x_\star) + f_\star.$$

The Gradient descent (GD) method is defined by the update rule

$$x_{t+1} = x_t - \gamma \nabla f(x_t) \tag{GD}$$

where γ is called step-size.

1. Prove (GD)'s iterates verify the relation $x_{t+1} - x_\star = (I - \gamma H)(x_t - x_\star)$.
2. Assuming H's eigenvalues λ verify $0 < \mu \leq \lambda \leq L$, provide the worst case rate of (GD). Propose a value for γ and provide the worst-case rate of (**??**) with this specific choice of step-size γ.
3. We now assume that $\mu = 0$ (i.e. H's eigenvalues can be arbitrarily small). The previous worst-case rate becomes 1. We then bound the function value. Prove that $f(x_t) - f_\star = \frac{1}{2}(x_0 - x_\star)^T H(I - \gamma H)^{2t}(x_0 - x_\star)$ and propose a worst case bound of $f(x_t) - f_\star$. Which γ would you consider ? What is the worst case value of $f(x_t) - f_\star$ with this specific step-size γ ?
4. We consider general first order methods of the form $x_{t+1} = x_0 - \sum_{s=0}^{t} \gamma_{t,s} \nabla f(x_s)$ with $\gamma_{t,t} \neq 0$. Prove that for each of those methods, there exists a sequence of polynomials $(P_t)_{t\in\mathbb{N}}$ with $deg(P_t) = t$ and $P_t(0) = 1$ such that $\forall t, x_t - x_\star = P_t(H)(x_0 - x_\star)$.
5. Prove the reciproqua of the previous question.
6. Assuming $\mu \neq 0$, prove the best method can be designed by solving

$$\begin{cases} \max\limits_{Q_t \in \mathbb{R}_t[X]} & Q_t\left(\frac{L+\mu}{L-\mu}\right) \\ \text{s.t.} & \sup_{X\in[-1,1]}(|Q_t(X)|) = 1 \end{cases} \tag{1}$$

   Hint : First consider the polynomial P_t associated with the considered method. Find the problem it solves, and use some translation and rescale of P_t to define Q_t.
7. Prove the t^{th} Tchebychev polynomial T_t, defined as verifying $\forall\theta,\ T_t(cos(\theta)) = cos(t\theta)$, is solution of this problem.
8. Find the associated method and convergence guarantee. Provide an equivalent of this rate.

   Hint : Tchebychev polynomials verify the recursion $T_{t+1} = 2XT_t - T_{t-1}$.

   Hint 2 : Tchebychev polynomials verify $\forall X > 1, T_t(X) \underset{t\to\infty}{\sim} \frac{1}{2}(X + \sqrt{X^2 - 1})^t$.
9. What is the stationary behavior of this method ? (Provide limits of the parameters). This method is called Heavy-ball (HB) method.

bonus : Find a convergence guarantee of HB.

10. We now assume $\mu = 0$. Write the corresponding problem over polynomials.
11. Propose some good candidate based on the Tchebychev family of polynomials.
12. Provide the associated method and convergence guarantee.