Problem 1

Consider the problem of minimizing

$$f(x) = -\ln(x) + x + 3, \quad x > 0,$$

using the Newton's method with the step-size $\gamma = 1$.

a) What is the minimizer x^* of f?

b) Find an interval $I \subseteq (0, \infty)$ such that for any $x^0 \in I$, the Newton's method converges to x^*.

Problem 2

Consider the following objective function

$$f(\boldsymbol{x}) = \mathrm{e}^{x_1^4 + x_2^4}.$$

a) Is f convex?

b) Is $\boldsymbol{g} = (-1, 1)$ a descent direction for $f(\boldsymbol{x})$ at $\boldsymbol{x} = (1, -1)$?

Problem 3

Consider the problem of minimizing the following function over $\mathbb{R}^2$

$$f(\boldsymbol{x}) = \frac{1}{2}x_1^2 + \frac{\theta}{2}x_2^2, \quad \theta > 0.$$

a) Compute the gradient and the Hessian of f.

b) What is the minimizer $\boldsymbol{x}^*$ of f? Is it unique?

c) Consider the gradient method with the step-size $\gamma = 2/(1+\theta)$. Express $f(\boldsymbol{x}^t)$ only in terms of θ and $f(\boldsymbol{x}^0)$. What can you conclude about the convergence rate of the method for this function?

Problem 4

Consider an objective function $f \in \mathcal{S}^1_{M,L}(\mathbb{R}^n)$ with a finite minimizer $f^* = f(\boldsymbol{x}^*)$ attained at $\boldsymbol{x}^* \in \mathbb{R}^n$. Then, for the step-size $\gamma = 1/L$, we saw in the class that the iterates of the gradient method satisfy

$$(f(\boldsymbol{x}^t) - f(\boldsymbol{x}^*)) \leq c^t \left[\frac{L}{2}\|\boldsymbol{x}^0 - \boldsymbol{x}^*\|^2\right] \quad \text{where} \quad c = 1 - \frac{M}{L}.$$

By using this result, prove the following bound that establishes the convergence of the iterates

$$\|\boldsymbol{x}^t - \boldsymbol{x}^*\|^2 \leq c^t \left[\frac{L}{M}\|\boldsymbol{x}^0 - \boldsymbol{x}^*\|^2\right].$$

Bonus Problem

Consider a quadratic function

$$f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^\mathsf{T}\boldsymbol{Q}\boldsymbol{x} - \boldsymbol{c}^\mathsf{T}\boldsymbol{x} + d \quad \text{with} \quad \boldsymbol{Q} \succ 0.$$

Show that if the initial point $\boldsymbol{x}^0$ is selected such that the vector $\boldsymbol{v} = \boldsymbol{x}^0 - \boldsymbol{x}^*$ is an eigenvector of $\boldsymbol{Q}$, then the gradient method can reach the optimal solution $\boldsymbol{x}^*$ in a single step, namely, $\boldsymbol{x}^1 = \boldsymbol{x}^*$.

Hint: Note that if $\boldsymbol{v} = \boldsymbol{x}^0 - \boldsymbol{x}^*$ is an eigenvector, then $\boldsymbol{Q}\boldsymbol{v} = \lambda\boldsymbol{v}$ where $\lambda > 0$ is an eigenvalue.