Problem 1

Consider the problem of minimizing

$$f(x) = 4x - \ln(x), \quad x > 0,$$

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

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

(b) Find an interval $I \in \mathbb{R}$ such that for any $x^0 \in I$, the Newton's method converges to x^*. Show that the order of convergence is at least quadratic.

Problem 2

Consider a set of $m+1$, possibly nonconvex, functions $\{f_0(\boldsymbol{x}), f_1(\boldsymbol{x}), \dots, f_m(\boldsymbol{x})\}$ defined over the whole $\mathbb{R}^n$ and the function $L : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ defined as

$$L(\boldsymbol{x}, \boldsymbol{\theta}) = f_0(\boldsymbol{x}) + \sum_{i=1}^{m} \theta_i f_i(\boldsymbol{x}).$$

Show that the following function over $\boldsymbol{\theta} \in \mathbb{R}^m$ is always convex

$$g(\boldsymbol{\theta}) = \sup_{\boldsymbol{x} \in \mathbb{R}^n} \{L(\boldsymbol{x}, \boldsymbol{\theta})\}.$$

Hint: You may rely on the definition of convexity and/or the properties of convex functions.

Problem 3

Consider the problem of minimizing the following function over $\mathbb{R}^3$

$$f(\boldsymbol{x}) = \frac{1}{2}\mathsf{exp}\left(\frac{1}{2}\boldsymbol{x}^\mathsf{T}\boldsymbol{Q}\boldsymbol{x} - \boldsymbol{c}^\mathsf{T}\boldsymbol{x} + d\right), \quad \text{where} \quad \boldsymbol{Q} = \begin{bmatrix} 2 & -1 & \theta \\ -1 & 2 & -1 \\ \theta & -1 & 2 \end{bmatrix}, \quad \boldsymbol{c} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \text{and} \quad d = 5.$$

(a) For what values of $\theta \in \mathbb{R}$ is the problem strictly convex?

(b) Is $\boldsymbol{g} = \boldsymbol{c} - \boldsymbol{Q}\boldsymbol{x}$ a descent direction for f at any $\boldsymbol{x} \in \mathbb{R}^3$?

(c) What is the minimizer $\boldsymbol{x}^*$ of f for $\theta = 0$?

Problem 4

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

$$f(\boldsymbol{x}) = \frac{1}{2}\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|^2 \quad \text{where} \quad \boldsymbol{A} \in \mathbb{R}^{m \times n} \quad \text{and} \quad \boldsymbol{b} \in \mathbb{R}^m.$$

(a) Provide the expressions for the gradient and the Hessian of f. Is f a convex function?

(b) Find the minimizer $\boldsymbol{x}^* \in \mathbb{R}^n$ of f, when $\boldsymbol{A}^\mathsf{T}\boldsymbol{A}$ is nonsingular.

(c) Assume that $\boldsymbol{A}^\mathsf{T}\boldsymbol{A}$ is nonsingular. Find a matrix $\boldsymbol{M} \in \mathbb{R}^{n \times n}$ such that for any $\boldsymbol{x}^0 \in \mathbb{R}^n$ the following single iteration

$$\boldsymbol{x}^1 = \boldsymbol{x}^0 - \boldsymbol{M}\nabla f(\boldsymbol{x}^0),$$

converges to the minimizer $\boldsymbol{x}^* \in \mathbb{R}^n$.

Hint: Note this result must be valid for any starting point $\boldsymbol{x}^0 \in \mathbb{R}^n$.

Bonus Problem

Consider a function $\boldsymbol{y} = \mathsf{G}(\boldsymbol{x})$ that takes in a vector $\boldsymbol{x} \in \mathbb{R}^n$ and produces another vector $\boldsymbol{y} \in \mathbb{R}^n$.

- We say that G is *Lipschitz continuous* with constant $L > 0$, when

$$\|\mathsf{G}(\boldsymbol{x}_1) - \mathsf{G}(\boldsymbol{x}_2)\| \leq L\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|, \quad \forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{R}^n.$$

  When $L < 1$, we say that G is *contractive*.

- We say that G is *strongly monotone* with constant $M > 0$, when

$$(\mathsf{G}(\boldsymbol{x}_1) - \mathsf{G}(\boldsymbol{x}_2))^{\mathsf{T}}(\boldsymbol{x}_1 - \boldsymbol{x}_2) \geq M\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|^2, \quad \forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{R}^n$$

(a) Show that for $f \in \mathcal{S}^1_{M,L}(\mathbb{R}^n)$, the gradient $\nabla f(\boldsymbol{x})$ is strongly monotone with constant M.

Hint: Use the definition of strong convexity at two different points $\boldsymbol{x}_1 \in \mathbb{R}^n$ and $\boldsymbol{x}_2 \in \mathbb{R}^n$.

(b) Show that for $f \in \mathcal{S}^1_{M,L}(\mathbb{R}^n)$ there exists $\gamma > 0$, such that the function $\mathsf{G}(\boldsymbol{x}) = \boldsymbol{x} - \gamma\nabla f(\boldsymbol{x})$ is contractive.

Hint: Use the fact that ∇f is both Lipschitz continuous and strongly monotone.

(c) Show that when $\mathsf{G}(\boldsymbol{x}) = \boldsymbol{x} - \gamma\nabla f(\boldsymbol{x})$ is contractive with a Lipschitz constant $0 < c < 1$, the iterates generated by the gradient method converge linearly to the minimizer $\boldsymbol{x}^* \in \mathbb{R}^n$ as follows

$$\|\boldsymbol{x}^t - \boldsymbol{x}^*\| \leq c^t\|\boldsymbol{x}^0 - \boldsymbol{x}^*\|, \quad \forall t \geq 1.$$