Problem 6.1: Convergence of Gradient Methods

Consider the quadratic function $f : \mathbb{R}^n \to \mathbb{R}$ defined as

$$f(x) = \frac{1}{2}x^\top Qx \tag{1}$$

where $Q \in \mathbb{R}^{n\times n}$ is a positive definite symmetric matrix. Suppose that $f(x)$ is minimized with a Gradient method using the update rule

$$x_{k+1} = x_k - \alpha_k \nabla f(x_k) \tag{2}$$

and exact line search where the stepsize α_k at iteration k is computed as the minimum α of

$$\theta(\alpha) = f(x_k - \alpha \nabla f(x_k)) \tag{3}$$

Let $\underline{\lambda}$ and $\overline{\lambda}$ be the minimum and maximum eigenvalues of the (Hessian) matrix Q, respectively. Show that for all iterations k, we have

$$f(x_{k+1}) \le \left(\frac{\overline{\lambda}-\underline{\lambda}}{\overline{\lambda}+\underline{\lambda}}\right)^2 f(x_k) \quad \text{or} \quad \frac{f(x_{k+1})}{f(x_k)} \le \left(\frac{\overline{\lambda}-\underline{\lambda}}{\overline{\lambda}+\underline{\lambda}}\right)^2 \tag{4}$$

Hint: For an arbitrary x_k calculate the next step according to the Gradient method. You also need to calculate the optimal step size α. Once you have done those, apply *Kantorovich inequality.* Let $Q \in \mathbb{R}^{n\times n}$ be a positive definite symmetric matrix. Then, for any vector $y \in \mathbb{R}^n$ with $y \neq 0$, we have

$$\frac{(y^\top y)^2}{(y^\top Qy)(y^\top Q^{-1}y)} \ge \frac{4\overline{\lambda}\,\underline{\lambda}}{(\overline{\lambda}+\underline{\lambda})^2} \tag{5}$$

where $\underline{\lambda}$ and $\overline{\lambda}$ are the minimum and maximum eigenvalues of Q, respectively.

Problem 6.2: Effect of Scaling on Gradient Method Convergence

Consider the following unconstrained optimization problem

$$\min_x .\ f(x) = (x_1 - 2)^2 + 5(x_2 + 6)^2 \tag{6}$$

where we denote the (quadratic) objective function $f : \mathbb{R}^2 \to \mathbb{R}$ of (6) as

$$f(x) = (x_1 - 2)^2 + 5(x_2 + 6)^2 \tag{7}$$

Suppose that we want to solve the problem (6) with a Gradient method using the update rule

$$x_{k+1} = x_k - \alpha_k \nabla f(x_k) \tag{8}$$

and exact line search where the stepsize α_k at iteration k is computed as the minimum α of

$$\theta(\alpha) = f(x_k - \alpha \nabla f(x_k)) \tag{9}$$

(a) Evaluate the convergence rate of the Gradient method applied to problem (6) with an arbitrary starting point. *Hint:* Use the the results of Exercise 6.1.

(b) Can you solve the problem faster by first modifying the objective function (7) and then applying the Gradient method to the modified problem? *Hint:* Try to find a variable substitution that gives the best convergence rate according to the results of Exercise 6.1.