

Project Work 1 is online from Monday 11.10.2021 <https://mycourses.aalto.fi/mod/assign/view.php?id=765881> and is due no later than **Friday 5.11.2021 23:55**. No homework this week.

Problem 6.1: Convergence of Gradient Methods

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

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

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

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

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

$$\theta(\alpha) = f(x_k - \alpha \nabla f(x_k)) \quad (3)$$

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

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

Problem 6.2: Effect of Scaling on Gradient Method Convergence

Consider the following unconstrained optimization problem

$$\min_x f(x) = (x_1 - 2)^2 + 5(x_2 + 6)^2 \quad (5)$$

where we denote the (quadratic) objective function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ of (5) as

$$f(x) = (x_1 - 2)^2 + 5(x_2 + 6)^2 \quad (6)$$

Suppose that we want to solve the problem (5) with a Gradient method using the update rule

$$x_{k+1} = x_k - \alpha_k \nabla f(x_k) \quad (7)$$

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

$$\theta(\alpha) = f(x_k - \alpha \nabla f(x_k)) \quad (8)$$

- (a) Evaluate the convergence rate of the Gradient method applied to problem (5) with an arbitrary starting point. *Hint:* Use the results of Exercise 6.1.
- (b) Can you solve the problem faster by first modifying the objective function (6) 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.