Optimization Algorithms
Weekly Exercise 2

Marc Toussaint
Learning & Intelligent Systems Lab, TU Berlin
Marchstr. 23, 10587 Berlin, Germany

Winter 2024/25

1 Gradient Descent

Consider the quadratic function $f(x) = x^\top A x + b^\top x + c$ with $A \in \mathbb{R}^{n\times n}$ symmetric and positive definite, $b \in \mathbb{R}^n$ and $c \in \mathbb{R}$. Starting at x_0, we do a line search using the gradient direction, i.e. $x' = x_0 - \alpha \nabla f(x_0)$. In this exercise, instead of doing backtracking, we can do exact line search, i.e., compute the optimal step length using the analytical expression of the quadratic function. Which is the best step size α?

2 Newton Step

a) Consider the quadratic function $f(x) = \frac{1}{2} x^\top A x + b^\top x + c$ with $A \in \mathbb{R}^{n\times n}$ symmetric and positive definite, $b \in \mathbb{R}^n$, $c \in \mathbb{R}$. Which x_{Min} minimizes $f(x)$? At a given location x, what is the Newton step δ that solves $\nabla^2 f(x)\delta = -\nabla f(x)$? How many iterations are required to reach x_{Min} if we take full Newton steps $x' = x + \delta$?

b) A fixed-stepsize Newton method iterates $x \leftarrow x + \alpha\delta$, for constant stepsize $\alpha \in [0, 1]$. Write down an explicit equation for the Newton iterates in the quadratic case of part a), i.e. find $x_k = \dots$ for the k-th iterate which only depends on A, b, c, α and x_0. For which values of α does it converge? How fast does it converge?

3 Levenberg-Marquardt regularization

Another small exercise to train you minimze a quadratic: Show that the regularized Newton step
$\delta = -(\nabla^2 f(x) + \lambda \mathbf{I})^{-1} \nabla f(x)$ minimizes

$$\min_\delta \left[\nabla f(x)^\top \delta + \frac{1}{2}\delta^\top \nabla^2 f(x) \delta + \frac{1}{2}\lambda \|\delta\|^2 \right] .$$