Exercise session notes - Week 8

Gradient Descent + Types of Convexity in Descent Methods

This week we started with the new chapter on unconstrained Optimization. The goal is to solve the following unconstrained problem

$$\min \quad f(x) \quad \text{for } x \in \text{dom}(f) \subseteq \mathbb{R}^n$$

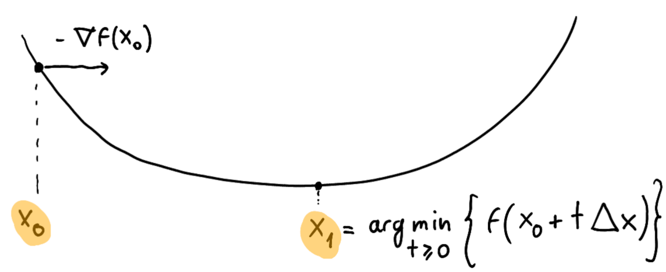
The only requirement is that the function f is twice differentiable.
We first note that this problem can be sometimes solved analytically: we find a point x^* with $\nabla f(x^*) = 0$, by solving a system of n equations. This implies that the point x^* is a local optimal solution, and if the function is convex, then x^* is also a global optimal solution. Often this method cannot be applied as it is much slow, for example if f does not have a closed form.

In such cases we apply the Descent Method: we produce a sequence of points $x_0, x_1, x_2, \ldots$ with

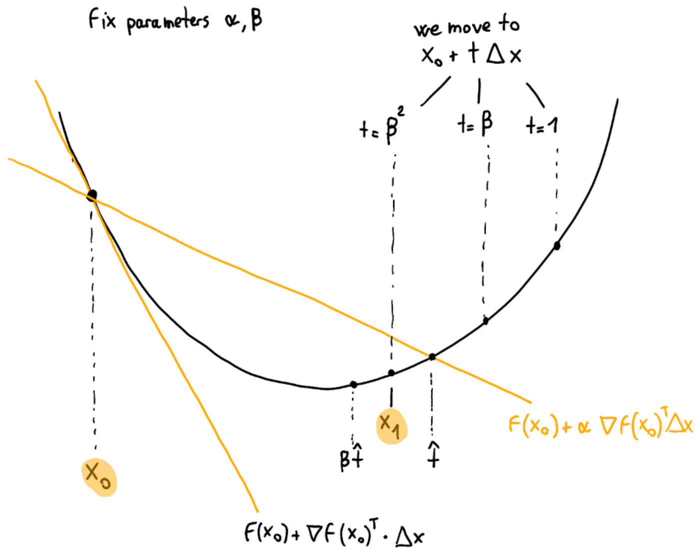
$$f(x_{j+1}) < f(x_j)\ ,\ f(x_j) \to f^*$$

We have to answer two main questions: Given a point x_0

- In which direction we move to get x_1? (Step Direction)
  In class we saw the Gradient descent, which uses $\Delta x = -\nabla f(x)$.
- How much we have to move in direction Δx? (Step Size)
  - Exact Line Search: we find the optimal step size in direction Δx



  - Backtracking Line Search: we move one unit in direction Δx and then backtrack until we meet some requirements



Finally we reviewed the different types of convexity:

- If f is not convex, then the Descent Method (usually) finds a local optimum; this is often the case in machine learning.
- If f is convex, then the Descent Method (usually) finds a global optimum, but not always (consider the convex function $-\log(x)$).
- If f is strictly convex, then the Descent Method (usually) finds a global optimum. This usually does not give us any additional properties compared to the convex case (consider the strictly convex function e^{-x})
- If f is stongly convex, then the Descent Method always finds a global optimum in a finite amount of time. This is the property that we usually look for.

We concluded the ex. class an important proposition on strong convexity that is a good practice exercise.

Proposition. If f is strongly convex with constant $m > 0$, then it is also coercive, i.e.

$$\lim_{\|x\| \to \infty} f(x) = \infty$$

In particular, there always exists an unique global minimum.