

WEEKLY PARTICIPATION 6: ITERATION COMPLEXITIES

Consider the problem

$$\operatorname{argmin}_{x \in \mathbb{R}} f(x),$$

where f is a convex function. We want to solve this optimization problem to ε -suboptimality in the function value, with $\varepsilon = 10^{-3}$. We have several options for optimization algorithms, let us see how many iterations they will each require.

- (1) If f is differentiable and β -smooth, how many iterations T of gradient descent suffice¹?
- (2) If f is differentiable and α -strongly convex, how many iterations of gradient descent suffice?
- (3) If f is differentiable and has convex condition number κ , how many iterations of gradient descent suffice?
- (4) If f is non-differentiable, how many iterations of subgradient descent suffice?
- (5) If f is twice-differentiable and we start sufficiently close to the optimizer x^* that the convergence theory for pure Newton phases apply, how many iterations of Newton's method suffice²? Assume the Hessian is L -Lipschitz and $\nabla^2 f(x^*) \geq \alpha I$.

For concreteness, take $\beta = 2$, $\alpha = 1$, define κ accordingly, and take $L = 1$.

¹For this and the other problems, give an expression for T that depends only on the given value of ε and the stated properties of the function. This means, e.g. ignoring the dependence on x_0 .

²Convince yourself that in the setting where the convergence theory for pure Newton phases applies and f is convex, the minimizer is unique.