Exercises: Gradient algorithms

Exercise 1 (A quadratic example in $\mathbb{R}^2$)**.** *Consider, for $\gamma > 0$, $f(x) = \frac{1}{2}(x_1^2 + \gamma x_2^2)$. We apply the gradient descent method with optimal step, starting at $x^{(0)} = (\gamma, 1)$.*

1. *Show that f is m-convex with M-Lipschitz gradient. Find the tightest m and M constants.*

2. *Show that*

$$x^{(k)} = \left(\gamma\Big(\frac{\gamma-1}{\gamma+1}\Big)^k, \Big(-\frac{\gamma-1}{\gamma+1}\Big)^k\right)$$

and

$$f(x^{(k)}) = \frac{\gamma(\gamma+1)}{2}\Big(\frac{\gamma-1}{\gamma+1}\Big)^{2k} f(x^{(0)})$$

3. *Show that, on this example, the convergence is exactly linear, that is $f(x^{(k)}) - v^\sharp$ is a geometric series. Give its reason. Compare with the theoretical bound.*

4. *When is this algorithm fast and slow ?*

Exercise 2 (Strongly convex - optimal step)**.** *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a m-convex $\mathcal{C}^2$ function. Define, for given $x^{(0)}$,*

$$\begin{aligned}
&\tilde{f}_k : t \mapsto f(x^{(k)} - t\nabla f(x^{(k)})) \\
&t^{(k)} = \operatorname*{arg\,min}_{t\in\mathbb{R}} \tilde{f}_k(t) \\
&x^{(k+1)} = x^{(k)} - t^{(k)}\nabla f(x^{(k)})
\end{aligned}$$

1. *Show that there exists $M \geq m$ such that $mI \preceq \nabla^2 f(x^{(k)}) \preceq MI$*

2. *Show that, for any interesting t (to be defined) we have*

$$\tilde{f}_k(t) \leq f(x^{(k)}) - t\|\nabla f(x^{(k)})\|_2^2 + \frac{Mt^2}{2}\|\nabla f(x^{(k)})\|_2^2$$

3. *Show that,*

$$f(x^{(k+1)}) \leq f(x^{(k)}) - \frac{1}{2M}\|\nabla f(x^{(k)})\|_2^2$$

4. *Show that $f(x^\sharp) \geq f(x) - \frac{1}{2m}\|\nabla f(x)\|_2^2$*

5. *Show that*

$$f(x^{(k+1)}) - f(x^\sharp) \leq (1 - \frac{m}{M})[f(x^{(k)}) - f(x^\sharp)]$$

6. *Show that the algorithm converges, and give its convergence speed.*

Exercise 3 (Strictly convex case)**.** *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a C^2, convex function with M-Lipschitz gradient such that $f(x^\sharp) = \inf f$. We define, for given $x^{(0)}$.*

$$x^{(k+1)} = x^{(k)} - t\nabla f(x^{(k)})$$

with $t \leq 1/M$.

1. *Show that, for all x and y $(y-x)^\top \nabla^2 f(z)(y-x) \leq M\|y-x\|_2^2$*

2. *Show that*

$$f(y) \leq f(x) + \nabla f(x)^\top (y-x) + \frac{M}{2}\|y-x\|^2$$

3. *Show that $f(x^{(k+1)}) \leq f(x^{(k)}) - \frac{t}{2}\|\nabla f(x^{(k)})\|^2$*

4. *Show that*

$$f(x^{(k+1)}) \leq v^\sharp + \nabla f(x^{(k)})^\top (x^{(k)} - x^\sharp) - \frac{t}{2}\|\nabla f(x^{(k)})\|^2$$

5. *Deduce that*

$$f(x^{(k+1)}) \leq v^\sharp + \frac{1}{2t}(\|x^{(k)} - x^\sharp\|^2 - \|x^{(k+1)} - x^\sharp\|^2)$$

6. *Show that*

$$\sum_{i=1}^{k} f(x^{(i)}) - v^\sharp \leq \frac{1}{2t}\|x^{(0)} - x^\sharp\|^2$$

7. *Conclude that*

$$f(x^{(k)}) - v^\sharp \leq \frac{1}{2kt}\|x^{(0)} - x^\sharp\|^2.$$