Exercises: differential calculus

Pierre Ablin

1 Gradient flows

We let $f : \mathbb{R}^p \to \mathbb{R}$ a differentiable function. Starting from $w^0 \in \mathbb{R}^p$, gradient descent with step-size $\eta > 0$ iterates

$$w^{k+1} = w^k - \eta \nabla f(w^k). \tag{1}$$

The convergence analysis seen in class may seem a bit unjustified. The behavior of such algorithm is more easily understood by looking at the gradient *flow*, which is the Ordinary Differential Equation (ODE), starting from $w(0) = w^0$:

$$\dot{w}(t) = -\nabla f(w(t)). \tag{2}$$

Indeed, Eq (1) is an Euler discretization of the gradient flow equation with step η, and as such we have $x^k \simeq x(\eta k)$.

1.1

We define $\phi(t) = f(w(t))$. Show that we have

$$\phi'(t) = -\|\nabla f(w(t))\|^2$$

1.2

We assume that f is bounded from below by f^*. Demonstrate that the function $t \to \|\nabla f(w(t))\|^2$ is integrable, and that

$$\inf_{t \leq T} \|\nabla f(w(t))\|^2 \leq \frac{f(w^0) - f^*}{T}.$$

1.3

Assume that f satisfies the Polyak-Lojasciewicz inequality (we recall that it is the case when f is strongly-convex):

$$f(w) - f^* \leq \frac{1}{2\mu}\|\nabla f(w)\|^2.$$

Demonstrate that $f(w(t))$ converges to f^*, and give the convergence rate.