Subgradient descent cannot converge for constant step size

April 3, 2018

Consider the problem

$$\min_{x \in \mathbb{R}} \quad f(x) := |x|$$

over scalars. The subdifferential of f is

$$\partial f(x) = \begin{cases} \{+1\} & x > 0 \\ \{-1\} & x < 0 \\ [-1, 1] & x = 0 \end{cases}$$

Fixed oracle Assume you have a subdifferential oracle, e.g. something that for x gives you $g_x \in \partial f(x)$ that works as follows:

$$g_x = \begin{cases} +1 & x \geq 0 \\ -1 & x < 0 \end{cases}$$

Now we run subgradient descent:

$$x^{(k+1)} = x^{(k)} - \alpha g_{x^{(k)}}$$

with starting point $x^{(0)} = \alpha$. Then we have

k	$x^{(k)}$	$g_{x^{(k)}}$
0	α	1
1	0	1
2	$-\alpha$	-1
3	0	1
4	$-\alpha$	-1
5	0	1

So even though we reach the optimum in 1 step, the oracle cannot tell us to terminate, and the algorithm will cycle forever.

If we start with $x^{(0)} = \alpha/2$ we will have

k	$x^{(k)}$	$g_{x^{(k)}}$
0	$\alpha/2$	1
1	$-\alpha/2$	-1
2	$\alpha/2$	1
3	$-\alpha/2$	-1
4	$\alpha/2$	1
5	$-\alpha/2$	-1

which never reaches the optimal solution.

Random oracle Can we do better if we force our oracle to sample uniformly from the subdifferential space? e.g.

$$g_x = \begin{cases} +1 & x > 0 \\ -1 & x < 0 \\ \text{Unif}[-1, 1] & x = 0 \end{cases}$$

In the case of $x^{(0)} = \alpha/2$, still no, because you will deterministically never hit $x^{(k)} = 0$. In the case of $x^{(0)} = \alpha$, maybe, since you might hit$x^{(k)} = 0$ sometimes (though it doesn't cycle, so not necessarily an infinite number of times). However, for random initialization, this will happen with probability 0.

Experiments with random oracle In both cases, $\alpha = 0.5$.

Figure on left: an evolution using random oracle, using two example starting points. Basically $x^{(k)} = 0$ never happens.

Figure on right: Histogram of best and final values using $x^{(0} \sim \text{Unif}([-1, 1])$. Numerically, it loos like

$$\Pr(|x_{best}| < \epsilon) = \frac{\alpha\epsilon}{2}, \quad \Pr(|x_{conv}| < \epsilon) = \alpha\epsilon$$

but converging to the optimal solution happens with probability 0