Exercise List: Convergence rates and complexity

Francis Bach and Robert M. Gower

February 10, 2019

1 Rate of convergence and complexity

All the algorithm we discuss in the course generate a sequence of random vectors x^t that converge to a desired x^* in some sense. Because the x^t's are random we always prove convergence in expectation. In particular, we focus on two forms of convergence, either showing that the difference of function values converges

$$\mathbb{E}\left[f(x^t) - f(x^*)\right] \longrightarrow 0,$$

or the expected norm difference of the iterates converges

$$\mathbb{E}\left[\|x^t - x^*\|^2\right] \longrightarrow 0.$$

Two important questions: 1) How fast is this convergence and 2) given an ϵ how many iterations t are needed before $\mathbb{E}\left[f(x^t) - f(x^*)\right] < \epsilon$ or $\mathbb{E}\left[\|x^t - x^*\|^2\right] < \epsilon$.

Ex. 1 — Consider a sequence $(\alpha_t)_t \in \mathbb{R}_+$ that converge to zero according to

$$\alpha_t \leq \frac{C}{t},$$

where $C > 0$. Given an $\epsilon > 0$, show that

$$t \geq \frac{C}{\epsilon} \quad \Rightarrow \alpha_t < \epsilon.$$

We refer to this result as a $O(1/\epsilon)$ iteration complexity.

Ex. 2 — Using that

$$\frac{1}{1-\rho}\log\left(\frac{1}{\rho}\right) \geq 1, \tag{1}$$

prove the following lemma.

Lemma 1.1. Consider the sequence $(\alpha_k)_k \in \mathbb{R}_+$ of positive scalars that converges to zero according to

$$\alpha_k \leq \rho^k \, \alpha_0, \tag{2}$$

where $\rho \in [0, 1)$. For a given $1 > \epsilon > 0$ we have that

$$k \geq \frac{1}{1-\rho} \log\left(\frac{1}{\epsilon}\right) \quad \Rightarrow \quad \alpha_k \leq \epsilon \, \alpha_0. \tag{3}$$

We refer to this as a $O(\log(1/\epsilon))$ iteration complexity.

Following the introduction, we can write $\alpha^t \stackrel{\text{def}}{=} \mathbb{E}\left[f(x^t) - f(x^*)\right]$ or $\alpha^t \stackrel{\text{def}}{=} \mathbb{E}\left[\|x^t - x^*\|^2\right]$. The type of convergence (2) is known as *linear convergence at a rate of* ρ^k.

Answer (Ex. 2) — *Proof.* First note that if $\rho = 0$ the result follows trivially. Assuming $\rho \in (0, 1)$, rearranging (2) and applying the logarithm to both sides gives

$$\log\left(\frac{\alpha_0}{\alpha_k}\right) \geq k \log\left(\frac{1}{\rho}\right). \tag{4}$$

Now using (1) and assuming that

$$k \geq \frac{1}{1-\rho} \log\left(\frac{1}{\epsilon}\right), \tag{5}$$

we have that

$$\begin{aligned} \log\left(\frac{\alpha_0}{\alpha_k}\right) &\stackrel{(4)}{\geq} k \log\left(\frac{1}{\rho}\right) \\ &\stackrel{(5)}{\geq} \frac{1}{1-\rho} \log\left(\frac{1}{\rho}\right) \log\left(\frac{1}{\epsilon}\right) \\ &\stackrel{(1)}{\geq} \log\left(\frac{1}{\epsilon}\right) \end{aligned}$$

Applying exponentials to the above inequality gives (3). □