Dhamma Kimpara and Sam Zhang

Problem 1

i) Yes, the cost is strongly convex. Write $f_t(x) = \frac{1}{2}\|Ax - b_t\|_2^2$, and $h_t(x) = Ax - b_t$. Then f_t is μ-strongly convex if the smallest eigenvalue of its Hessian $\nabla^2 f_t(x) = A$ is positive, with μ equal to that eigenvalue. By construction, then, $\mu = 1/\sqrt{\kappa}$.

ii) We know from class that for convergence, we require the step size α to be in the range $(0, 2/L]$. Since L is the Lipschitz constant of $\nabla f = Ax - b_t$, we have that since

$$\|Ax - b_t - (Ay - b_t)\| = \|Ax - Ay\| \leq \|A\| \|x - y\|,$$

that $L \leq \|A\| = 1$, since the largest singular value of A is 1. Thus we must pick $\alpha \in (0, 2]$. The bound we derived in class uses

$$\rho = \max\{|1 - \alpha\mu|, |1 - \alpha L|\} = \max\{\left|1 - \alpha/\sqrt{\kappa}\right|, |1 - \alpha|\}.$$

This gives the finite bound

$$\|x_t - x_t^*\| \leq \rho^t \|x_0 - x_0^*\| + \sigma \sum_{i=0}^{t} \rho^i = \rho^t \|x_0 - x_0^*\| + \frac{1 - \rho^t}{1 - \rho}$$

with the asymptotic result

$$\limsup_{t\to\infty} \|x_t - x_t^*\| \leq \frac{1}{1 - \rho}.$$

With this specific problem, since $\kappa = 100$, we have that for an arbitrary pick of $\alpha = 1$ that $rho = 99/100$, so the asymptotic bound is $1/(1 - \rho) = 100$. As seen in Figure 1, the tracking error does not appear to typically approach either the finite or asymptotic bounds here.

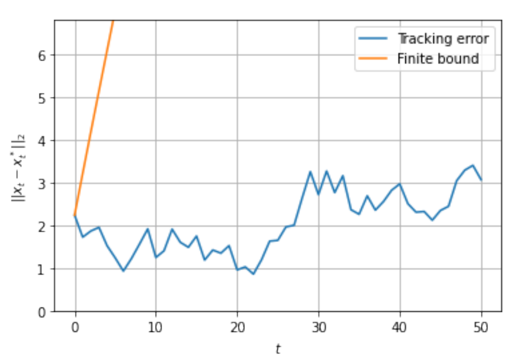

Figure 1: Tracking error of one run of online gradient descent vs. finite bounds