Dhamma Kimpara and Sam Zhang
ECEN 5008-005 – Homework #4

Problem 1

i) Yes, the cost is strongly convex. Write $g(x) = \frac{1}{2}\|x\|_2^2$, and $h_t(x) = Ax - b_t$. Then g is $\mu = 1$-strongly convex, and h_t is affine for all t, so our cost function is the composition $f = g \circ h_t$, which is also 1-strongly convex.

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 is in terms of

$$\rho = \max\{|1 - \alpha\mu|, |1 - \alpha L|\} = |1 - \alpha|$$

since $\mu = L = 1$ for our problem. Then the bound at step t is

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

with the asymptotic result

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

Thus we have a tradeoff: if we pick α to be near 1, then ρ is small, and our asymptotic limit is good, but the initial error $\|x_0 - x_0^*\|$ affects our estimate for longer. If we pick α to be near 0, then the initial error decays rapidly, but the asymptotic error is controlled very weakly.