Convex Programming
Problem Set 5 – CS 6515/4540 (Fall 2025)

Solutions to problem statement 5, Q20

20 Gradient Descent for Strongly-Convex Functions

By α-strong convexity,

$$f(x_t) - f(x^\star) \leq \langle \nabla f(x_t),\, x_t - x^\star \rangle - \frac{\alpha}{2}\|x_t - x^\star\|^2. \tag{1}$$

Rewriting the inner product using the GD-step update:

$$\langle \nabla f(x_t),\, x_t - x^\star \rangle = \frac{1}{\eta_t}\langle x_t - x_{t+1},\, x_t - x^\star \rangle = \frac{1}{2\eta_t}\Big(\|x_t - x^\star\|^2 - \|x_{t+1} - x^\star\|^2 + \|x_t - x_{t+1}\|^2\Big).$$

Using $\|x_t - x_{t+1}\|^2 = \eta_t^2\|\nabla f(x_t)\|^2 \leq \eta_t^2 G^2$ gives

$$\begin{aligned} f(x_t) - f(x^\star) &\leq \frac{1}{2\eta_t}\big(\|x_t - x^\star\|^2 - \|x_{t+1} - x^\star\|^2\big) + \frac{\eta_t}{2}G^2 - \frac{\alpha}{2}\|x_t - x^\star\|^2 \\ &= \Big(\frac{1}{2\eta_t} - \frac{\alpha}{2}\Big)\|x_t - x^\star\|^2 - \frac{1}{2\eta_t}\|x_{t+1} - x^\star\|^2 + \frac{\eta_t}{2}G^2. \end{aligned} \tag{2}$$

Setting $\eta_t = \dfrac{1}{\alpha(t+1)}$ from the hint. We get,

$$\frac{1}{2\eta_t} - \frac{\alpha}{2} = \frac{\alpha(t+1)}{2} - \frac{\alpha}{2} = \frac{\alpha t}{2}, \qquad \frac{1}{2\eta_t} = \frac{\alpha(t+1)}{2},$$

so (2) becomes

$$f(x_t) - f(x^\star) \leq \frac{\alpha}{2}\Big(t\|x_t - x^\star\|^2 - (t+1)\|x_{t+1} - x^\star\|^2\Big) + \frac{G^2}{2}\eta_t. \tag{3}$$

Summing over $t = 1, \ldots, T$

$$\sum_{t=1}^{T}\big(t\|x_t - x^\star\|^2 - (t+1)\|x_{t+1} - x^\star\|^2\big) = \|x_1 - x^\star\|^2 - (T+1)\|x_{T+1} - x^\star\|^2 \leq \|x_1 - x^\star\|^2.$$

Hence from (3) we obtain

$$\sum_{t=1}^{T}\big(f(x_t) - f(x^\star)\big) \leq \frac{\alpha}{2}\|x_1 - x^\star\|^2 + \frac{G^2}{2}\sum_{t=1}^{T}\eta_t. \tag{4}$$

Strong convexity implies a relationship between gradient norm and distance to the minimizer:

$$\langle \nabla f(x), x - x^\star \rangle \geq \alpha\|x - x^\star\|^2 \quad (\text{since } \nabla f(x^\star) = 0).$$

By Cauchy–Schwarz, $\|\nabla f(x)\|\,\|x - x^\star\| \geq \alpha\|x - x^\star\|^2$, when $\|x - x^\star\| \leq \|\nabla f(x)\|/\alpha$. Using $\|\nabla f(x)\| \leq G$ we get

$$\|x_1 - x^\star\| \leq \frac{G}{\alpha} \quad\Longrightarrow\quad \frac{\alpha}{2}\|x_1 - x^\star\|^2 \leq \frac{G^2}{2\alpha}.$$

Since $\eta_t = \dfrac{1}{\alpha(t+1)}$,

$$\sum_{t=1}^{T} \eta_t = \frac{1}{\alpha} \sum_{t=1}^{T} \frac{1}{t+1} = \frac{1}{\alpha}\big(H_{T+1} - 1\big) \le \frac{1 + \ln T}{\alpha},$$

Plugging into (4) gives us

$$\sum_{t=1}^{T} \big(f(x_t) - f(x^\star)\big) \le \frac{G^2}{2\alpha} + \frac{G^2}{2} \cdot \frac{1 + \ln T}{\alpha} = \frac{G^2}{2\alpha}\big(1 + (1 + \ln T)\big).$$

$$\sum_{t=1}^{T} \big(f(x_t) - f(x^\star)\big) \le \frac{G^2}{2\alpha}(1 + \ln T).$$

Therefore, by convexity of f,

$$f\Big(\frac{1}{T}\sum_{t=1}^{T} x_t\Big) \le \frac{1}{T}\sum_{t=1}^{T} f(x_t),$$

and so

$$f\Big(\frac{1}{T}\sum_{t=1}^{T} x_t\Big) - f(x^\star) \le \frac{1}{T}\sum_{t=1}^{T} \big(f(x_t) - f(x^\star)\big) \le \frac{G^2}{2\alpha T}(1 + \ln T).$$

Hence, we have the desired bound

$$f\Big(\frac{1}{T}\sum_{t=1}^{T} x_t\Big) - f(x^\star) \le \frac{G^2(1 + \ln T)}{2\alpha T}.$$