Online Convex Optimization & Complexity
Problem Set 6 – CS 6515/4540 (Fall 2025)

Answers to HW6, Q21.

22 Online Gradient Descent

Define the potential

$$\Phi(t) = \frac{\|x_t - x^*\|^2}{2\eta}.$$

By nonexpansiveness of projection,

$$2\eta\Phi(t+1) = \|x_{t+1} - x^*\|^2 \le \|x_t - \eta G_t - x^*\|^2 = \|x_t - x^*\|^2 - 2\eta G_t^\top (x_t - x^*) + \eta^2 \|G_t\|^2.$$

$$\Phi(t+1) - \Phi(t) \le -G_t^\top (x_t - x^*) + \frac{\eta}{2}\|G_t\|^2.$$

$$G_t^\top (x_t - x^*) \le \Phi(t) - \Phi(t+1) + \frac{\eta}{2}\|G_t\|^2.$$

By convexity,

$$f_t(x_t) - f_t(x^*) \le \nabla f_t(x_t)^\top (x_t - x^*).$$

Taking expectation and using $\mathbb{E}[G_t] = \nabla f_t(x_t)$ gives

$$\mathbb{E}\big[f_t(x_t) - f_t(x^*)\big] \le \mathbb{E}\big[G_t^\top (x_t - x^*)\big].$$

Combining with the previous inequality, taking expectation, and summing $t = 1$ to T:

$$\begin{aligned}
\mathbb{E}\Big[\sum_{t=1}^{T} f_t(x_t) - f_t(x^*)\Big] &\le \mathbb{E}\Big[\sum_{t=1}^{T} \big(\Phi(t) - \Phi(t+1)\big)\Big] + \frac{\eta}{2}\sum_{t=1}^{T} \mathbb{E}\big[\|G_t\|^2\big] \\
&= \mathbb{E}[\Phi(1) - \Phi(T+1)] + \frac{\eta}{2}\sum_{t=1}^{T} \mathbb{E}\big[\|G_t\|^2\big].
\end{aligned}$$

Since $\Phi(T+1) \ge 0$ and $\|x_1 - x^*\| \le D$,

$$\mathbb{E}[\Phi(1) - \Phi(T+1)] \le \Phi(1) = \frac{\|x_1 - x^*\|^2}{2\eta} \le \frac{D^2}{2\eta}.$$

Also $\mathbb{E}[\|G_t\|^2] \le G^2$, so

$$\frac{\eta}{2}\sum_{t=1}^{T} \mathbb{E}\big[\|G_t\|^2\big] \le \frac{\eta G^2 T}{2}.$$

Thus

$$\mathbb{E}\Big[\sum_{t=1}^{T} f_t(x_t) - f_t(x^*)\Big] \le \frac{D^2}{2\eta} + \frac{\eta G^2 T}{2}.$$

Setting $\eta = \frac{D}{G\sqrt{T}}$. We get

$$\boxed{\mathbb{E}\Big[\sum_{t=1}^{T} f_t(x_t) - f_t(x^*)\Big] \le DG\sqrt{T}}.$$