1. **Strength in Convexity.** A function $f : \mathbb{R}^n \to \mathbb{R}$ is called *α-strongly-convex* with respect to norm $\|\cdot\|$ if
$$f(\mathbf{y}) \geq f(\mathbf{x}) + \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{\alpha}{2}\|\mathbf{y} - \mathbf{x}\|^2.$$
I.e., if the function is not just convex, but "locally it grows at least as fast as a quadratic".

For this problems, focus on just the Euclidean norm $\|\cdot\|_2$, and assume that $\|\nabla f(x)\| \leq G$.

(a) Modify the basic gradient descent analysis to show that the update rule $\mathbf{x}_{t+1} \leftarrow \mathbf{x}_t - \eta_t \nabla f(\mathbf{x}_t)$ with suitably chosen η_t can be used to find $\widehat{\mathbf{x}} \in \mathbb{R}^n$ with
$$f(\widehat{\mathbf{x}}) - f(\mathbf{x}^*) \leq O\left(\frac{G^2 \log T}{\alpha \cdot T}\right).$$
Observe: the assumption of strong convexity gives better convergence guarantees (i.e., the dependence on T is better, and there is no dependence on $D = \|\mathbf{x}_0 - \mathbf{x}^*\|$. Check that your proof also works for the constrained case, for some convex body $K \subseteq \mathbb{R}^n$.

(b) Show that this analysis also works in the online case to give a regret bound, if each function is strongly convex. (A slightly harder problem is to remove the $\log T$ term in the numerator in the offline case. Why does this improvement not extend to the online case?)

(c) Also, show that if x^* is the minimizer of the function with the convex body K, then $f(x) \leq f(x^*) + \varepsilon$ implies that $\|x - x^*\|^2 \leq O(\alpha\varepsilon)$.

2. **That's the Norm.** We define a differentiable convex function $f : K \to \mathbb{R}$ to be G-Lipschitz with respect to norm $\|\cdot\|$ if
$$\frac{|f(x) - f(y)|}{\|x - y\|} \leq G.$$
Show that this is equivalent to $\|\nabla f(x)\|_* \leq G$ for all $x \in K$. [1]

Similarly, show that f being α-strongly-convex with respect to $\|\cdot\|$ is equivalent to $\|\nabla f(x) - \nabla f(y)\|_* \geq \alpha\|x - y\|$. And that f being β-smooth with respect to $\|\cdot\|$ is equivalent to $\|\nabla f(x) - \nabla f(y)\|_* \leq \beta\|x - y\|$.

3. **These are "Small" Numbers.** For an integer k, define $\langle k \rangle = 1 + \lceil \log_2(|k| + 1) \rceil$; for a rational p/q (with p, q coprime, $q > 0$), define $\langle p/q \rangle = \langle p \rangle + \langle q \rangle$; for a matrix $R = (r_{ij})$ of rationals, define $\langle M \rangle = \sum_{i,j} \langle r_{ij} \rangle$. Let $det(R)$ denote the determinant of R.

(a) If R is an $n \times n$ matrix, show that $\langle det(R) \rangle \leq poly(n, \langle R \rangle)$.

Now consider the LP $\min\{c^T x \mid Ax \geq b\}$, where $x \in \mathbb{R}^n$, and $A \in \mathbb{R}^{m \times n}$ with $m \geq n$.

(b) If each of the numbers in A and b are rationals having size at most S, and if x^* is a basic feasible solution, give an upper bound on the size of each entry of x^*. In particular, show that each entry x_i^* is a rational number with size at most $K = \text{poly}(n \cdot S)$.

[1] Given norm $\|\cdot\|$, the dual norm is defined as $\|\theta\|_* = \max\{\langle \theta, x \rangle \mid \|x\| = 1\}$.

(c) If the LP also has a finite optimum, and the numbers in c have size at most S, infer that the optimal value of the LP has size $O((K+S)n)$.

Hint: Google for 'Cramer's rule' and 'Hadamard Inequality'.

4. **The Oracle Separates.** Given a graph with edge costs c_e, the following LP is called the Held-Karp/subtour elimination LP, and is a relaxation for the traveling salesman problem. Give a separation oracle for it.

$$\begin{aligned} &\min \textstyle\sum_e c_e x_e \\ &\textstyle\sum_{e \in \partial S} x_e \geq 2 \qquad \forall \varnothing \subsetneq S \subsetneq V \\ &0 \leq x_e \leq 1. \end{aligned}$$

Here's another LP with variables $\{x_i\}_{i=1}^n$, non-negative costs $\{c_i\}_{i=1}^n$, and sizes $\{s_{ij}\}_{i,j=1}^n$: again give a separation oracle.

$$\begin{aligned} &\min \textstyle\sum_i c_i x_i \\ &\textstyle\sum_{i \in S} x_i s_{i,|S|} \geq |S| \qquad \forall S \subseteq [n] \\ &0 \leq x_i \leq 1. \end{aligned}$$