1 Mastery set [25 points] (Aaditya)

Lemma 1: If $f : D \to \mathbb{R}$ is convex and $m \in \mathbb{R}$, then $U_f(m) := \{x \in D : f(x) \leq m\}$ is convex.

Proof: For $x, y \in U_f(m), \theta \in [0, 1]$, if $z = \theta x + (1-\theta)y$, then

$$f(z) \leq \theta f(x) + (1-\theta) f(y) \leq m,$$

so $z \in U_f(m)$. ■

A [2 + 2] If $S = \emptyset$, then S is trivially convex. Otherwise, for $x \in S$, $S = U_f(f(x))$ is convex, by Lemma 1. If f is strictly convex and $\exists x, y \in S$ with $x \neq y$, for $\theta \in (0,1)$, $z = \theta x + (1-\theta)y$,

$$f(z) < \theta f(x) + (1-\theta) f(y) = f(x),$$

contradicting the fact that x minimizes f.

B [2 + 2] If A is orthogonal, $A^T A = I$, whose only singular value is 1. For

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

both A and B are orthogonal, but $\boxed{\frac{1}{2}A + \frac{1}{2}B}$ is not orthogonal.

C [2 + 2 + 2] By Lemma 1, the unit ball $U_{\|\cdot\|}(1)$ is convex. Since, by part B, all orthogonal matrices are in the unit ball, it follows that any convex combination of orthogonal matrices is in the unit ball. By the triangle inequality, for any norm $\|\cdot\|$,

$$\begin{aligned} \|x\| &= \|x - y + y\| \leq \|x - y\| + \|y\| \\ \text{and } \|y\| &= \|y - x + x\| \leq \|x - y\| + \|x\|, \end{aligned}$$

which implies

$$-\|x - y\| \leq \|x\| - \|y\| \leq \|x - y\|,$$

and so $|\|x\| - \|y\|| \leq \|x - y\|$ (this statement is precisely the Reverse Triangle Inequality. Hence, any norm is Lipschitz with respect to itself, with Lipschitz constant $\boxed{L = 1.}$ ■

D [5] We already showed that the convex hull of orthogonal matrices in contained in the unit ball of $\|\cdot\|_{\text{op}}$.

If A satisfies $\|A\|_{\text{op}} \leq 1$, then A has a singular value decomposition $A = U\Sigma V^T$, where U and V are orthogonal and Σ is diagonal and has entries in $[-1, 1]$. Let $\mathcal{D}$ denote the set of diagonal matrices with diagonal entries in $\{-1, 1\}$. It is geometrically obvious that $\text{conv}(\{-1,1\}^n) = [-1,1]^n$, and hence $\Sigma \in \text{conv}(\mathcal{D})$. Also, any $D \in \mathcal{D}$ is clearly orthogonal. Thus, since $\Sigma = \sum_{D \in \mathcal{D}} \theta_D D$, where $\sum_{D \in \mathcal{D}} \theta_D = 1$,

$$A = U \left(\sum_{D \in \mathcal{D}} \theta_D D \right) V^T = \sum_{D \in \mathcal{D}} \theta_D U D V^T$$

is in the convex hull of orthogonal matrices (as any product of orthogonal matrices is itself orthogonal). ■

E [4+2] Rearranging the first-order definition of strong convexity,

$$\begin{aligned} f(y) - \langle \nabla f(x), y - x \rangle &\geq f(x) + \frac{\lambda}{2}\|y - x\|_2^2 \\ f(x) + \langle \nabla f(y), y - x \rangle &\geq f(y) + \frac{\lambda}{2}\|y - x\|_2^2. \end{aligned}$$

Adding these inequalities and cancelling terms,

$$\langle \nabla f(y) - \nabla f(x), y - x \rangle \geq \lambda \|y - x\|_2^2.$$

Then, by Cauchy-Schwarz,

$$\begin{aligned} \|\nabla f(y) - \nabla f(x)\| &\geq \frac{\langle \nabla f(y) - \nabla f(x), y - x \rangle}{\|y - x\|_2} \\ &\geq +\lambda \|y - x\|_2. \quad ■ \end{aligned}$$

If ∇f is Lipschitz with Lipschitz constant L, then

$$L\|y - x\|_2 \geq \|\nabla f(y) - \nabla f(x)\| \geq +\lambda \|y - x\|_2,$$

so $\boxed{L \geq \lambda.}$