SC 646 HW#1 Due 10th February 2026

1. **[Primer on Strong Convexity]** Let f be a strongly-convex continuously differentiable function with modulus $\mu > 0$. Show that

$$f(y) \le f(x) + \nabla f(x)^\intercal (y-x) + \frac{1}{2\mu}\|\nabla f(y) - \nabla f(x)\|^2, \quad \forall x, y$$

[Hint: Consider the function $\phi_x(z) := f(z) - \nabla f(x)^\intercal z$. Show that $\phi_x(\cdot)$ is strongly convex with same modulus μ, and then apply PL-inequality.]

2. [**AGD vs HB**]: Consider the function $f(x) := \log\left(x + \sqrt{1+x^2}\right) + x^2$, for $x \in \mathbb{R}$. Is the function strongly-convex? Does the function admit Lipschitz-gradient? Implement the accelerated gradient method (AGD) and heavy-ball (HB) methods in Python, and compare their performances for this test functions in terms of iterations required for convergence. Share your implementations.

3. [**Contraction Coefficient**]: Consider the simplest first-order gradient descent scheme

$$x_{k+1} = x_k - \alpha \nabla f(x_k).$$

We consider gradient descent as a fixed-point iteration:

$$x_{k+1} \leftarrow G_\alpha(x_k).$$

The gradient step G_α is a *contraction* if:

$$\|G_\alpha(x) - G_\alpha(y)\|_2 \le L_G \|x - y\|_2 \quad \text{for some } L_G < 1.$$

The idea is that since the optimum $x^\star$ is a fixed-point of G_α, we would want the fixed-point iteration $x_{k+1} \leftarrow G_\alpha(x_k)$ to converge to $x^\star$. Below, we need to show three key results that let us analyze Polyak's Heavy Ball method or the Accelerated Gradient Descent method.

   (a) If the gradient step $G_\alpha(x)$ is a contraction, then the Gradient Descent converges linearly to $x^\star$.

   (b) Let f be twice continuously differentiable, L-smooth and μ-strongly convex, then the contraction coefficient L_G is at most $\max\{|1-\alpha\mu|, |1-\alpha L|\}$. [Hint: You may need to use Mean-Value Theorem]

   (c) For steepest descent, we need to choose the step-size in order to minimize the contraction coefficient L_G. Find the optimal step size and show that

$$\|x_{k+1} - x^\star\|_2 \le \left(\frac{\kappa - 1}{\kappa + 1}\right)^{k+1} \|x_0 - x^\star\|_2,$$

   where $\kappa := \dfrac{L}{\mu}$ is the condition number of the problem.

4. [**Kernel SVMs**] We briefly discussed kernel-SVMs in class. This particular problem requires some hands-on experience with Python (scikit-learn). You are required to submit a Jupyter notebook that consists of following features:

   (a) Generate $n = 100$ points uniformly inside a unit circle. Assign them labels $y = +1$

   (b) Randomly generate $n = 100$ points uniformly inside a disc bounded between $x_1^2 + x_2^2 = 4$ and $x_1^2 + x_2^2 = 6$. Assign them labels $y = -1$.

   (c) Run a simple linear SVM and report the performance (Show decision boundaries).

   (d) Run a kernel-SVM and report the performance (Show decision boundaries). A 'poly' (polynomial) kernel should suffice. You may try other kernels, such as 'rbf' (radial basis functions).

5. **[Linear Convergence]** Consider the following optimization problem,

$$\underset{x \in \mathbb{R}^d}{\arg\min} \quad F(x) = f(x) + g(x),$$

where ∇f is L-Lipschitz but g may be non-smooth. $F(\cdot)$ is said to satisfy the proximal-PL inequality if

$$\mathcal{D}_g(x, L) \geq 2\mu(F(x) - F^*),$$

where $\mathcal{D}_g(x, \alpha) \coloneqq -2\alpha \min_y \{\langle \nabla f(x), y - x \rangle + \alpha \|y - x\|^2 + g(y) - g(x)\}$. Show that F converges linearly, and find the corresponding contraction coefficient.

6. **[Necessity and Sufficiency of KKT Conditions]** Show that for a convex optimization problem with strong duality, $x^\star$ and $(\lambda^\star, \nu^\star)$ are primal and dual optimal solutions, respectively *if and only if* $x^\star$, $(\lambda^\star, \nu^\star)$ satisfy KKT conditions.

7. **[KKT conditions in action]** Solve the following optimization problem:

$$\begin{aligned} \text{maximize} \quad & xy \\ \text{subject to,} \quad & x + y^2 \leq 2 \\ & x, y \geq 0. \end{aligned}$$

8. **[Condition for Convexity]** Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function such that $\mathrm{dom}\ f$ is convex and for all $x, y \in \mathrm{dom}\ f$, $f(y) \geq f(x) + \nabla f(x)^\top (y - x)$. Show that f is a convex function.