Sample midterm questions

March 11, 2026

1. (100 points) Let $X \subset \mathbb{R}^n$ be nonempty. We say that a sequence $(x_k)_{k\in\mathbb{N}}$ in $\mathbb{R}^n$ is *quasi-Fejér monotone* with respect to X if there exists a summable positive sequence $(\epsilon_k)_{k\in\mathbb{N}}$ such that

$$\|x_{k+1} - x\|^2 \leqslant \|x_k - x\|^2 + \epsilon_k, \quad \forall x \in X, \quad \forall k \in \mathbb{N}.$$

   (a) (15 points) Prove that if a sequence $(x_k)_{k\in\mathbb{N}}$ in $\mathbb{R}^n$ is quasi-Fejér monotone with respect to X, then the sequence is bounded, i.e., there exists $M > 0$ such that $\|x_k\| \leqslant M$ for all $k \in \mathbb{N}$.

   (b) (20 points) Prove that if a sequence $(x_k)_{k\in\mathbb{N}}$ in $\mathbb{R}^n$ is quasi-Fejér monotone with respect to X and admits a subsequence converging to $x^* \in X$, then $x_k \to x^*$.

   Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function whose infimum is obtained at $X \subset \mathbb{R}^n$. Assume that f is L-Lipschitz, that is, $|f(x) - f(y)| \leqslant L\|x - y\|$ for all $x, y \in \mathbb{R}^n$. Denote by ∂f the convex subdifferential of f.

   (c) (15 points) Prove that $\|v\| \leqslant L$ for all $v \in \partial f(x)$, $x \in \mathbb{R}^n$.

   We next study the iterates of the subgradient method for minimizing f. Let $(x_k)_{k\in\mathbb{N}}$ be a sequence generated by $x_0 \in \mathbb{R}^n$ and $x_{k+1} \in x_k - \alpha_k \partial f(x_k)$, where $(\alpha_k)_{k\in\mathbb{N}}$ is a positive sequence.

   (d) (25 points) Derive a condition on $(\alpha_k)_{k\in\mathbb{N}}$ under which $(x_k)_{k\in\mathbb{N}}$ is quasi-Fejér monotone with respect to X.

   (e) (25 points) Derive a condition on $(\alpha_k)_{k\in\mathbb{N}}$ under which $(x_k)_{k\in\mathbb{N}}$ converges to a minimizer of f.

2. (100 points) Let f be a convex and L-smooth function. Let $g : \mathbb{R}^n \to \mathbb{R}$ to be a continuously differentiable and σ-strongly convex function. Denote $D_g(x, y)$ as the Bregman divergence induced by g.

   (a) (25 points) Show that

$$D_g(x, y) + D_g(y, z) = D_g(x, z) + \langle \nabla g(z) - \nabla g(y), x - y\rangle.$$

   (b) (25 points) Consider the update rule $\nabla g(x_{k+1}) = \nabla g(x_k) - \alpha_k \nabla f(x_k)$. Show that for all $y \in \mathbb{R}^n$,

$$f(y) - f(x_k) \geq \frac{1}{\alpha_k}(D_g(y, x_{k+1}) - D_g(x_k, x_{k+1}) - D_g(y, x_k)).$$

   (c) (30 points) By selecting $\alpha_k = \sigma/L$, it holds that

$$D_g(x_k, x_{k+1}) \leq \alpha_k(f(x_k) - f(x_{k+1})).$$

   (d) (20 points) Show that for all $T \in \mathbb{N}^+$,

$$\min_{k=0,\dots,T-1} f(x_k) \leq f(x^*) + \frac{LD_g(x^*, x_0)}{\sigma T}.$$

3. (100 points) Let $f := \sum_{i=1}^n f_i$, where $f_i : \mathbb{R}^n \to \mathbb{R}$ is convex and L-smooth for all $i = 1, \dots, n$. Assume f is μ-strongly convex. Consider the following algorithm:

$$x_{k+1} = x_k - \alpha_k \nabla f_{i_k}(x_k), \quad i_k \sim \mathrm{Unif}[n], \quad \forall\, k \in \mathbb{N}.$$

   Define $\sigma_f^* := \inf_{x^* \in \arg\min f} \mathbb{E}[\|\nabla f_i(x^*) - \nabla f(x^*)\|^2]$.

(a) (25 points) Show that for every $x^* \in \arg\min f$, we have $\sigma_f^* = \mathbb{E}[\|\nabla f_i(x^*) - \nabla f(x^*)\|^2]$.

(b) (20 points) Show that

$$\mathbb{E}[\|\nabla f_i(x)\|^2] \leq 4L(f(x) - \inf f) + 2\sigma_f^*.$$

(c) (30 points) With step size $\alpha_k = \alpha \in (0, 1/(2L)]$, show that for all $k \geq 0$,

$$\mathbb{E}[\|x_k - x^*\|^2] \leq (1 - \alpha\mu)^k \|x_0 - x^*\|^2 + \frac{2\alpha}{\mu}\sigma_f^*.$$

(d) (25 points) Show that for every $\epsilon > 0$, we can guarantee that $\mathbb{E}[\|x_T - x^*\|^2] \leq \epsilon$ provided that

$$\alpha := \min\left\{\frac{\mu\epsilon}{4\sigma_f^*}, \frac{1}{2L}\right\}, \quad \text{and} \quad T \geq \max\left\{\frac{4\sigma_f^*}{\epsilon\mu^2}, \frac{2L}{\mu}\right\} \log\left(\frac{2\|x_0 - x^*\|^2}{\epsilon}\right).$$