TTIC 31070: Convex Optimization
Homework 4

Professor Zhiyuan Li

Spring 2026

Problem 1 (Heavy-ball momentum on quadratics). *Let*

$$f(x) = \frac{1}{2}x^\top Ax - b^\top x,$$

where $A \in \mathbb{R}^{d\times d}$ is symmetric positive definite and $\mu I \preceq A \preceq LI$. Let $x^\star = A^{-1}b$ and $\kappa := L/\mu$. Consider the heavy-ball iteration

$$x_{t+1} = x_t - \alpha\nabla f(x_t) + \beta(x_t - x_{t-1}), \qquad t \geq 1.$$

(a) *Let $e_t := x_t - x^\star$. Derive the error recurrence*

$$e_{t+1} = (I - \alpha A + \beta I)e_t - \beta e_{t-1}.$$

Then diagonalize A and show that, in an eigen-direction with eigenvalue $\lambda \in [\mu, L]$, the scalar error satisfies the characteristic equation

$$r^2 - (1 + \beta - \alpha\lambda)r + \beta = 0.$$

(b) *For fixed α, β, let*

$$\rho_\lambda(\alpha, \beta)$$

denote the largest modulus of the two roots of

$$r^2 - (1 + \beta - \alpha\lambda)r + \beta = 0.$$

Find one choice of $\alpha > 0$ and $\beta \in [0, 1)$ and a universal constant $c_0 > 0$ such that

$$\sup_{\lambda\in[\mu,L]} \rho_\lambda(\alpha, \beta) \leq 1 - \frac{c_0}{\sqrt{\kappa}}.$$

Conclude that heavy-ball on strongly convex quadratics has spectral factor

$$1 - \Theta(1/\sqrt{\kappa}),$$

or equivalently an $O(\sqrt{\kappa}\log(1/\varepsilon))$ iteration scale. Hint: *You may use the following Schur condition: the roots of*

$$r^2 - cr + \beta = 0$$

have modulus at most ρ if and only if

$$\beta \leq \rho^2, \qquad |c| \leq \rho + \frac{\beta}{\rho}.$$

Try $\alpha = 1/L$ and $\beta = (1 - 1/\sqrt{\kappa})^2$. It is enough to prove the bound with $c_0 = 1/2$.

(c) *Compare this factor with the best fixed-stepsize gradient-descent factor on strongly convex quadratics. Why does this spectral proof not prove the general smooth strongly convex acceleration theorem?*

Problem 2 (Quadratic cone constraints, SOCP barriers, and Newton steps)**.** *Lecture 17 already derived the robust-LP-to-SOCP modeling step. Here we start directly from quadratic constraints that are naturally represented by second-order cone blocks, and focus on the barrier Hessian seen by Newton's method.*

Let $c, a_i \in \mathbb{R}^n$, $b_i \in \mathbb{R}$, $A_i \in \mathbb{R}^{k_i \times n}$, and $p_i \in \mathbb{R}^{k_i}$. Consider

$$\min_x c^\top x \qquad s.t. \qquad \|A_i x + p_i\|_2 \le a_i^\top x + b_i, \quad i = 1, \dots, m.$$

Let

$$\mathcal{Q}^{k+1} := \{(\tau, z) \in \mathbb{R} \times \mathbb{R}^k : \tau \ge \|z\|_2\}$$

be the second-order cone. We write the constraints directly as affine cone blocks

$$z_i(x) := (a_i^\top x + b_i,\ A_i x + p_i) \in \mathcal{Q}^{k_i+1}.$$

Equivalently, this is the quadratic inequality

$$\|A_i x + p_i\|_2^2 \le (a_i^\top x + b_i)^2$$

together with $a_i^\top x + b_i \ge 0$.

(a) *For a cone variable $w = (\tau, z) \in \operatorname{int} \mathcal{Q}^{k+1}$, use the logarithmic barrier*

$$\Phi(\tau, z) := -\log(\tau^2 - \|z\|_2^2).$$

Write the fixed-t barrier objective as a function of x, and state its strict-feasibility domain.

(b) *Compute $\nabla\Phi(\tau, z)$ and $\nabla^2\Phi(\tau, z)$. You may use the compact notation*

$$J := \begin{bmatrix} 1 & 0 \\ 0 & -I_k \end{bmatrix}, \qquad w = (\tau, z), \qquad \delta(w) := w^\top J w = \tau^2 - \|z\|_2^2 .$$

(c) *Write the affine cone block as*

$$z_i(x) := (a_i^\top x + b_i,\ A_i x + p_i) = B_i x + d_i.$$

Let F_t denote the barrier objective from part (a)*. Using the chain rule, write $\nabla F_t(x)$ and $\nabla^2 F_t(x)$. At a strictly feasible point where $\nabla^2 F_t(x) \succ 0$, write the Newton direction.*

(d) *Suppose the current Newton direction is Δx. What condition must a line-search step size $\alpha > 0$ satisfy to keep the next iterate strictly feasible? Briefly explain how this conic barrier formulation differs from treating*

$$\|A_i x + p_i\|_2^2 - (a_i^\top x + b_i)^2 \le 0$$

as an ordinary scalar inequality constraint.

Problem 3 (Log-loss portfolios and mirror-map geometry)**.** *A portfolio is a vector $x \in \Delta_d$, where*

$$\Delta_d := \{x \in \mathbb{R}_+^d : \sum_{i=1}^d x_i = 1\}.$$

The coordinate x_i is the fraction of current wealth invested in asset i. At round t, the coordinate $r_{t,i}$ is the one-period price relative of asset i: for example, $r_{t,i} = 1.05$ means that asset i grows by 5% during the round. If the learner enters round t with wealth W_t and chooses portfolio x_t, then

$$W_{t+1} = W_t \sum_{i=1}^d x_{t,i} r_{t,i} = W_t \langle r_t, x_t \rangle .$$

Thus

$$W_{T+1} = W_1 \prod_{t=1}^T \langle r_t, x_t \rangle .$$

Taking logarithms turns this multiplicative wealth objective into an additive online objective,

$$\log W_{T+1} = \log W_1 + \sum_{t=1}^T \log \langle r_t, x_t \rangle ,$$

so maximizing final wealth is equivalent to minimizing cumulative log loss.

At round t, the learner chooses $x_t \in \mathrm{ri}(\Delta_d)$, then observes a nonzero return vector $r_t \in \mathbb{R}_+^d \setminus \{0\}$, and suffers

$$\ell_t(x_t) := -\log \langle r_t, x_t \rangle .$$

For $r \in \mathbb{R}_+^d \setminus \{0\}$, write

$$\ell_r(x) := -\log \langle r, x \rangle .$$

We compare three mirror maps:

$$\Phi_2(x) = \frac{1}{2} \|x\|_2^2 , \qquad \Phi_{\mathrm{ent}}(x) = \sum_{i=1}^d x_i \log x_i, \qquad \Phi_{\log}(x) = -\sum_{i=1}^d \log x_i.$$

Initialize $x_1 = \frac{1}{d}\mathbf{1}$.

(a) *Compute $\nabla \ell_r(x)$. For a twice differentiable mirror map Φ, define*

$$\|a\|_{\Phi,x,*}^2 := a^\top (\nabla^2 \Phi(x))^{-1} a.$$

Compute

$$\|\nabla \ell_r(x)\|_{\Phi_2,x,*}^2 , \qquad \|\nabla \ell_r(x)\|_{\Phi_{\mathrm{ent}},x,*}^2 , \qquad \|\nabla \ell_r(x)\|_{\Phi_{\log},x,*}^2 .$$

Which of these three quantities is uniformly bounded over $x \in \mathrm{ri}(\Delta_d)$ and $r \in \mathbb{R}_+^d \setminus \{0\}$?

(b) *For $u \in \Delta_d$ and $0 < \rho < 1$, define the smoothed comparator*

$$u^\rho := (1-\rho)u + \rho \frac{1}{d}\mathbf{1}.$$

Prove $u_i^\rho \geq \rho/d$ and

$$\sum_{t=1}^T \ell_t(u^\rho) \leq \sum_{t=1}^T \ell_t(u) + T \log \frac{1}{1-\rho}.$$

Also prove the potential bounds

$$D_{\Phi_2}(u^\rho, x_1) \le 1, \qquad D_{\Phi_{\mathrm{ent}}}(u^\rho, x_1) \le \log d, \qquad D_{\Phi_{\log}}(u^\rho, x_1) \le d \log \frac{1}{\rho}.$$

(c) *Let*

$$\Delta_d^\rho := \{x \in \Delta_d : x_i \ge \rho/d \ \forall i\}.$$

Write explicitly the Euclidean projected-gradient update on Δ_d^ρ*, the entropy mirror-descent update on* Δ_d^ρ*, and the log-barrier mirror-descent update on* $\mathrm{ri}(\Delta_d)$*. In all three cases use* $g_t = \nabla \ell_t(x_t)$*.*

(d) *Define the local decrement of a mirror step at* x *against return* r *by*

$$M_\Phi^X(x, r; \eta) := \sup_{y \in X} \left\{ \eta \left\langle -\frac{r}{\langle r, x\rangle}, x - y \right\rangle - D_\Phi(y, x) \right\}.$$

Prove the following three local bounds. For Euclidean geometry on Δ_d^ρ*,*

$$M_{\Phi_2}^{\Delta_d^\rho}(x, r; \eta) \le \frac{\eta^2 d^2}{2\rho^2}.$$

For entropy geometry on Δ_d^ρ*, if* $\eta \le \rho/d$*, then*

$$M_{\Phi_{\mathrm{ent}}}^{\Delta_d^\rho}(x, r; \eta) \le \frac{\eta^2 d}{\rho}.$$

For log-barrier geometry on $\mathrm{ri}(\Delta_d)$*, if* $0 < \eta \le 1/2$*, then*

$$M_{\Phi_{\log}}^{\mathrm{ri}(\Delta_d)}(x, r; \eta) \le \eta^2.$$

For the first two bounds, it may help to first prove

$$\left\| \frac{r}{\langle r, x\rangle} \right\|_2 \le \frac{d}{\rho}, \qquad \sum_{i=1}^d x_i \left(\frac{r_i}{\langle r, x\rangle} \right)^2 \le \frac{d}{\rho} \qquad (x \in \Delta_d^\rho).$$

For the log-barrier bound, you may use the scalar identity

$$\sup_{q > 0} \{\alpha(q - 1) + \log q - q + 1\} = -\alpha - \log(1 - \alpha), \qquad 0 \le \alpha < 1.$$

(e) *You may use the standard mirror-descent inequality*

$$\sum_{t=1}^T (\ell_t(x_t) - \ell_t(u)) \le \frac{D_\Phi(u, x_1)}{\eta} + \frac{1}{\eta} \sum_{t=1}^T M_t,$$

where M_t *is the local Bregman decrement of the corresponding mirror step. Using parts* (b) *and* (d)*, first prove the intermediate bounds*

$$\mathrm{Reg}_T^{(2)}(u) \le \frac{1}{\eta} + \frac{\eta T d^2}{2\rho^2} + T \log \frac{1}{1 - \rho},$$

$$\mathrm{Reg}_T^{(\mathrm{ent})}(u) \le \frac{\log d}{\eta} + \frac{\eta T d}{\rho} + T \log \frac{1}{1-\rho},$$

and

$$\mathrm{Reg}_T^{(\log)}(u) \le \frac{d \log(1/\rho)}{\eta} + \eta T + T \log \frac{1}{1-\rho}.$$

Then choose η and ρ to prove, for every $u \in \Delta_d$, the regret estimates

$$\mathrm{Reg}_T^{(2)}(u) = O\left(\sqrt{d}\, T^{3/4}\right),$$

$$\mathrm{Reg}_T^{(\mathrm{ent})}(u) = O\left((d \log d)^{1/3} T^{2/3}\right),$$

and

$$\mathrm{Reg}_T^{(\log)}(u) = O\left(\sqrt{dT \log T} + d \log T\right).$$

Briefly explain the geometric reason the log-barrier bound avoids the truncation-dependent local cost.