CS395T: Continuous Algorithms
Homework III

Kevin Tian

Due date: March 13, 2025, start of class (3:30 PM).

Please list all collaborators on the first page of your solutions. Unless we have discussed and I have specified otherwise, homework is not accepted if it is not turned in by hand at the start of class, or turned in electronically on Canvas by then. Send me an email to discuss any exceptions.

1 Problem 1

Let $\mathbf{A} \in \mathbb{R}^{n \times d}$ have rows $\{\mathbf{a}_i\}_{i \in [n]} \in \mathbb{R}^d$, and denote its leverage scores by $\{\boldsymbol{\tau}_i(\mathbf{A})\}_{i \in [n]}$.

(i) Prove that for all $i \in [n]$,

$$\boldsymbol{\tau}_i(\mathbf{A}) = \min_{\substack{\mathbf{x} \in \mathbb{R}^n \\ \mathbf{A}^\top \mathbf{x} = \mathbf{a}_i}} \|\mathbf{x}\|_2^2 .$$

(ii) Prove that there is a projection matrix[1] $\mathbf{P} \in \mathbb{S}_{\succeq \mathbf{0}}^{n \times n}$ such that $\boldsymbol{\tau}_i(\mathbf{A}) = \mathbf{P}_{ii}$, for all $i \in [n]$.

2 Problem 2

Let $\mathbf{v} \in \mathbb{R}^d$ be a random vector such that $\mathbb{E}\mathbf{v} = \mathbf{0}_d$, $\mathbb{E}\mathbf{v}\mathbf{v}^\top = \mathbf{I}_d$, and for some $C \geq 1$,[2]

$$\mathbb{E} \langle \mathbf{v}, \mathbf{u} \rangle^4 \leq C \|\mathbf{u}\|_2^4 , \text{ for all } \mathbf{u} \in \mathbb{R}^d.$$

(i) Prove that[3] $d^2 \leq \mathbb{E} \|\mathbf{v}\|_2^4 \leq Cd^2$.

(ii) Prove that for any $\mathbf{X} \in \mathbb{S}_{\succeq \mathbf{0}}^{d \times d}$ with $\mathrm{Tr}(\mathbf{X}) = 1$, $\mathbb{E}[\|\mathbf{v}\|_2^2 \langle \mathbf{v}\mathbf{v}^\top, \mathbf{X} \rangle] \leq Cd$.

3 Problem 3

Let $R \geq 0$ and $\mathbf{v} \in \mathbb{R}^d_{\geq 0}$. Give an algorithm which runs in time $O(d \log d)$, and computes[4]

$$\mathrm{argmin}_{\|\mathbf{x}\|_1 \leq R} \left\{ \frac{1}{2} \|\mathbf{x} - \mathbf{v}\|_2^2 \right\}.$$

4 Problem 4

For $\mathbf{K} \in \mathbb{S}_{\succ \mathbf{0}}^{d \times d}$, define the quantity

$$\kappa^\star_{\mathrm{diag}}(\mathbf{K}) := \min_{\substack{\mathbf{W} \in \mathbb{S}_{\succ \mathbf{0}}^{d \times d} \\ \mathbf{W} \text{ is diagonal}}} \frac{\boldsymbol{\lambda}_1\left(\mathbf{W}^{-\frac{1}{2}} \mathbf{K} \mathbf{W}^{-\frac{1}{2}}\right)}{\boldsymbol{\lambda}_d\left(\mathbf{W}^{-\frac{1}{2}} \mathbf{K} \mathbf{W}^{-\frac{1}{2}}\right)}.$$

[1] Recall from Section 2.2, Part VI that $\mathbf{P} \in \mathbb{S}_{\succeq \mathbf{0}}^{n \times n}$ is an orthogonal projection matrix iff $\mathbf{P}^2 = \mathbf{P}$.

[2] This condition is a special case of the vector $\mathbf{v}$ satisfying a condition called 2*-to-*4*-hypercontractivity*. The purpose of this question is to formalize a simple scenario where $\mathbb{E}\|\mathbf{G}\|_{\mathrm{op}}^2$ can be much larger than $\mathbb{E}\|\mathbf{G}\|_{\mathrm{op}} \langle \mathbf{G}, \mathbf{X} \rangle$ for $\mathbf{X} \in \mathbb{S}_{\succeq \mathbf{0}}^{d \times d}$ with $\mathrm{Tr}(\mathbf{X}) = 1$, by taking $\mathbf{G} = \mathbf{v}\mathbf{v}^\top$, as sketched at the end of Section 2, Part VIII.

[3] The scalar inequality $a^2 b^2 \leq \frac{1}{2}(a^4 + b^4)$ may be helpful to bound any cross terms.

[4] Note that this is one step in the projected gradient descent algorithm in Section 3, Part XII. It may be helpful to consider the Karush-Kuhn-Tucker conditions for this constrained optimization problem.

(i) Prove that $\kappa^\star_{\text{diag}}(\mathbf{K}) \leq \kappa$ iff there is a diagonal matrix $\mathbf{W} \in \mathbb{S}^{d\times d}_{\succ \mathbf{0}}$ such that

$$\mathbf{W} \preceq \mathbf{K} \preceq \kappa \mathbf{W}.$$

(ii) Prove that if $\mathbf{D} \in \mathbb{S}^{d\times d}_{\succ \mathbf{0}}$ is a diagonal matrix,

$$\kappa^\star_{\text{diag}}(\mathbf{K} + \mathbf{D}) \leq \kappa^\star_{\text{diag}}(\mathbf{K}).$$

5 Problem 5

Let $f : \mathcal{X} \to \mathbb{R}$ be a convex self-concordant barrier, where $\mathcal{X} \subseteq \mathbb{R}^d$, and assume that f is of Legendre type. Prove that f^* is also self-concordant.[5] You can use without proof that if $\mathbf{M}(t)$ takes $t \in \mathbb{R}$ to invertible matrices,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left((\mathbf{M}(t))^{-1}\right) = -\left(\mathbf{M}(t)\right)^{-1}\left(\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{M}(t)\right)\left(\mathbf{M}(t)\right)^{-1}.$$

[5] The characterization of convex conjugates immediately following Lemma 3, Part III may be helpful.