SI251 - Convex Optimization Quiz

2024-03-11 13:00:00

1. You can use Word, Latex or handwriting to complete this assignment. If you want to submit a handwritten version, scan it clearly.
2. The **report** has to be submitted as a PDF file to Blackboard, other formats are not accepted.
3. The submitted file name is **student_id+your_student_name.pdf**.
4. Late policy: You have 4 free late days for the quarter and may use up to 2 late days per assignment with no penalty. Once you have exhausted your free late days, we will deduct a late penalty of 25% per additional late day. Note: The timeout period is recorded in days, even if you delay for 1 minute, it will still be counted as a 1 late day.
5. You are required to follow ShanghaiTech's academic honesty policies. You are not allowed to copy materials from other students or from online or published resources. Violating academic honesty can result in serious sanctions.

Any plagiarism will get Zero point.

1. (30 pt) **Dual cone of psd.** What is the dual cone of $\mathbb{S}_+^n$? Please prove that. Hint: The matrix inner product is $< A, B >= \mathbf{tr}(\mathrm{AB}^{\mathrm{T}})$.

   Solution:

   The dual cone of $\mathbb{S}_+^n$ is $\mathbb{S}_+^n$ itself. In one hand, recall the definition of dual cone of : $K^* = \left\{\boldsymbol{y} \mid \boldsymbol{y}^T\boldsymbol{x} \geq 0 \text{ for all } \boldsymbol{x} \in K\right\}$. In the other hand, for any $\boldsymbol{X}, \boldsymbol{Y} \in \mathbb{S}^n$, we have

   $$< \boldsymbol{X}, \boldsymbol{Y} > \mathbf{tr}(XY^T) = \mathbf{tr}(XY) = \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i^{\boldsymbol{X}} \lambda_j^{\boldsymbol{Y}} \left\langle v_i^{\boldsymbol{X}}, v_j^{\boldsymbol{Y}} \right\rangle^2$$

   where $\lambda_i^{\boldsymbol{X}}$ is the the i-th eigenvalue of X corresponding to the (normalized) eigenvector $v_i^{\boldsymbol{X}}$, which are orthogonal to each other (and equivalently for Y).

   For any $\boldsymbol{Y} \in \mathbb{S}_+^n$, if $\boldsymbol{X}$ is positive semi-definite (i.e., all $\lambda_i^{\boldsymbol{X}} \geq 0$), $< \boldsymbol{X}, \boldsymbol{Y} > \geq 0$. Conversely, if $\boldsymbol{X}$ is not psd (i.e., there exists $\lambda_i^{\boldsymbol{X}} < 0$), there exists $\boldsymbol{Y}$ that makes $< \boldsymbol{X}, \boldsymbol{Y} >< 0$. For example, we may construct matrix $\boldsymbol{Y}$ where $v_i^{\boldsymbol{X}} = v_i^{\boldsymbol{X}}$ and $\lambda_i^{\boldsymbol{Y}} = \begin{cases} 0 & \lambda_i^{\boldsymbol{X}} \geq 0 \\ 1 & \textit{otherwise} \end{cases}$.

2. (30 pt) **Convexity of hyperbolic sets.** Please show that $\left\{x \in \mathbb{R}_+^n \mid \prod_{i=1}^n x_i \geq 1\right\}$ is a convex set. Hint: If $a, b \geq 0$ and $0 \leq \theta \leq 1$, then $a^\theta b^{1-\theta} \leq \theta a + (1-\theta)b$.

   Solution:

   Assume that $\prod_i x_i \geq 1$ and $\prod_i y_i \geq 1$. Using the inequality in the hint, we have

   $$\prod_i (\theta x_i + (1-\theta)y_i) \geq \prod x_i^\theta y_i^{1-\theta} = \left(\prod_i x_i\right)^\theta \left(\prod_i y_i\right)^{1-\theta} \geq 1.$$

3. (40 pt) **Eigenvalue functions.** Please prove:

   (a) $f(\boldsymbol{X}) = \lambda_{\max}(\boldsymbol{X}) = \sup_{\boldsymbol{y}\neq \boldsymbol{0}} \frac{\boldsymbol{y}^T\boldsymbol{X}\boldsymbol{y}}{\boldsymbol{y}^T\boldsymbol{y}}$ is convex on $\mathbb{S}^n$.

   (b) $f(\boldsymbol{X}) = \mathbf{tr}(\boldsymbol{X}) = \lambda_1(\boldsymbol{X}) + \cdots + \lambda_n(\boldsymbol{X})$ is linear on $\mathbb{S}^n$.

   (c) $f(\boldsymbol{X}) = \mathbf{tr}(\boldsymbol{X}^{-1}) = \frac{1}{\lambda_1(\boldsymbol{X})} + \cdots + \frac{1}{\lambda_n(\boldsymbol{X})}$ is convex on $\mathbb{S}_{++}^n$.

   Solution:

   (a) Define $\boldsymbol{X}_1, \boldsymbol{X}_2 \in \mathbb{S}^n$. Then we have

   $$\begin{aligned} \theta\lambda_{\max}(\boldsymbol{X}_1) + (1-\theta)\lambda_{\max}(\boldsymbol{X}_2) &= \theta \sup_{\boldsymbol{y}\neq\boldsymbol{0}} \frac{\boldsymbol{y}^T\boldsymbol{X_1}\boldsymbol{y}}{\boldsymbol{y}^T\boldsymbol{y}} + (1-\theta)\sup_{\boldsymbol{y}\neq\boldsymbol{0}} \frac{\boldsymbol{y}^T\boldsymbol{X_2}\boldsymbol{y}}{\boldsymbol{y}^T\boldsymbol{y}} \\ &\geq \sup_{\boldsymbol{y}\neq\boldsymbol{0}} \frac{\theta\boldsymbol{y}^T\boldsymbol{X_1}\boldsymbol{y}}{\boldsymbol{y}^T\boldsymbol{y}} + \frac{(1-\theta)\boldsymbol{y}^T\boldsymbol{X_2}\boldsymbol{y}}{\boldsymbol{y}^T\boldsymbol{y}} \\ &= \lambda_{\max}(\theta\boldsymbol{X}_1 + (1-\theta)\boldsymbol{X}_2) \end{aligned}$$

   (b) Define $\boldsymbol{X}_1, \boldsymbol{X}_2 \in \mathbb{S}^n$. Obviously, we have

   $$\mathbf{tr}(\theta\boldsymbol{X}_1 + (1-\theta)\boldsymbol{X}_2) = \theta\,\mathbf{tr}(\boldsymbol{X}_1) + (1-\theta)\,\mathbf{tr}(\boldsymbol{X}_1)$$

(c) Define $g(t) = f(Z + tV)$, where $Z \succ 0$ and $V \in \mathbf{S}^n$.

$$\begin{aligned}
g(t) &= \mathrm{tr}\left((Z+tV)^{-1}\right) \\
&= \mathrm{tr}\left(Z^{-1}\left(I + tZ^{-1/2}VZ^{-1/2}\right)^{-1}\right) \\
&= \mathrm{tr}\left(Z^{-1}Q(I+t\Lambda)^{-1}Q^T\right) \\
&= \mathrm{tr}\left(Q^TZ^{-1}Q(I+t\Lambda)^{-1}\right) \\
&= \sum_{i=1}^{n}\left(Q^TZ^{-1}Q\right)_{ii}(1+t\lambda_i)^{-1},
\end{aligned}$$

where we used the eigenvalue decomposition $Z^{-1/2}VZ^{-1/2} = Q\Lambda Q^T$. In the last equality, we express g as a positive weighted sum of convex functions $1/(1+t\lambda_i)$, hence it is convex.