SI251 - Convex Optimization homework 3

Deadline: 2022-12-13 23:59:59

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 Gradescope, other formats are not accepted.
3. You have to write your assignment in English, otherwise you will get a 50% penalty of your score.
4. The submitted file name is **student_id+your_student_name.pdf**.
5. 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.
6. 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. (45 pts) **Proximal Algorithms**

   (a) Compute the proximal operator $\mathbf{prox}_f$ of convex function $f(\mathbf{X}) = \lambda\|\mathbf{X}\|_*$, where $\mathbf{X} \in \mathbb{R}^{\mathbf{d}\times\mathbf{m}}$ is a matrix and $\lambda \in \mathbb{R}_+$ is the regularization parameter.

   (b) For $\mathbf{0} \prec \boldsymbol{\lambda}, \boldsymbol{\mu} \prec \mathbf{1}$, define $f_{\boldsymbol{\lambda},\boldsymbol{\mu}} : \mathbb{R}^n \to \mathbb{R}$ as

   $$f_{\boldsymbol{\lambda},\boldsymbol{\mu}}(\boldsymbol{x}) = \boldsymbol{\lambda}^T \boldsymbol{x}_+ + \boldsymbol{\mu}^T \boldsymbol{x}_-,$$

   where $\boldsymbol{x}_+ = \max\{\boldsymbol{x}, \mathbf{0}\}$ and $\boldsymbol{x}_- = \max\{-\boldsymbol{x}, \mathbf{0}\}$, the maximum is taken elementwise. Give a simple expression for the proximal operator of $\mathbf{prox}_{f_{\boldsymbol{\lambda},\boldsymbol{\mu}}}$.

   (c) On the page 18 of the slides for Lecture 11, please prove the orthogonal mapping: if $f(\mathbf{x}) = g(\mathbf{Q}\mathbf{x})$ with $\mathbf{Q}$ orthogonal ($\mathbf{Q}\mathbf{Q}^\mathbf{T} = \mathbf{Q}^\mathbf{T}\mathbf{Q}) = \mathbf{I}$), then

   $$\mathbf{prox}_f(\mathbf{x}) = \mathbf{Q}^T \mathbf{prox}_g(\mathbf{Q}\mathbf{x})$$

2. (25 pts) **Smoothing for Nonsmooth Optimization**

   Moreau envelope of a convex function f with parameter $\mu > 0$ is defined as

   $$M_{\mu f}(\boldsymbol{x}) := \inf_{\boldsymbol{z}} \left\{ f(\boldsymbol{z}) + \frac{1}{2\mu}\|\boldsymbol{x} - \boldsymbol{z}\|_2^2 \right\}.$$

   (a) Show that $M_{\mu f}$ is convex.

   (b) Show that $M_{\mu f}$ is $\frac{1}{\mu}$-smooth.

3. (30 pts) **Alternating Direction Method of Multipliers**

   For the sparse Gaussian graphical model given below, provide the ADMM update (the derivation process is required) for each variable at the t-th iteration

   $$\begin{aligned} &\min_{\Theta} - \log\det \boldsymbol{\Theta} + \langle \boldsymbol{\Theta}, \boldsymbol{S} \rangle + \mathbb{I}_{\mathbf{S}_+}(\boldsymbol{\Theta}) + \lambda\|\boldsymbol{\Psi}\|_1 \\ &\text{s.t.} \;\; \boldsymbol{\Theta} = \boldsymbol{\Psi}, \end{aligned}$$

   where $\mathbb{I}_C(\cdot)$ is the indicator function associated with the set C and $\mathbb{S}_+ := \{\boldsymbol{X} \mid \boldsymbol{X} \succeq \mathbf{0}\}$.