SI251 - Convex Optimization, Spring 2021
Final Exam

Note: We are interested in the reasoning underlying the solution, as opposed to simply the answer. Thus, solutions with the correct answer but without adequate explanation will not receive full credit; on the other hand, partial solutions with explanation will receive partial credit. Within a given problem, you can assume the results of previous parts in proving later parts (e.g., it is fine to solve part 3) first, assuming the results of parts 1) and 2)). Your use of resources should be limited to printed lecture slides, lecture notes, homework, homework solutions, general resources, class reading and textbooks, and other related textbooks on optimization. You should not discuss the final exam problems with anyone or use any electronic devices. Detected violations of this policy will be processed according to ShanghaiTech's code of academic integrity. Please hand in the exam papers and answer sheets at the end of exam.

I. Basic Knowledge

1. A second-order cone is $K = \{(\boldsymbol{x}, t) \in \mathbb{R}^{n+1} | \|\boldsymbol{x}\|_2 \leq t\}$ and the dual cone of a cone K is $K^* = \{\boldsymbol{y} | \boldsymbol{y}^\mathsf{T}\boldsymbol{x} \geq 0 \text{ for all } \boldsymbol{x} \in K\}$. Determine whether second-order cones are self dual or not? Give necessary explanations. (5 points)
2. Find all of the stationary points of $f(x, y) = 100(y - x^2)^2 - x^2$ on $\mathbb{R}^2$. For each stationary point, determine if it is a local minimum, local maximum, or neither. Justify your answer. (5 points)
3. Consider a convex function $f : \mathbb{R}^d \to \mathbb{R}$, which is subdifferentiable at each $\boldsymbol{x} \in \mathbb{R}^d$. Let $\boldsymbol{x}^*$ be a minimizer of f. Show that $\mathbf{0} \in \partial f(\boldsymbol{x}^*)$ if and only if $\boldsymbol{x}^* \in \arg\min_{\boldsymbol{x} \in \mathbb{R}^d} f(\boldsymbol{x})$. (5 points)
4. Compute the proximal operator prox_f for $f(\boldsymbol{X}) = \lambda \|\boldsymbol{X}\|_*$, where $\|\cdot\|_*$ is nuclear norm, $\boldsymbol{X} \in \mathbb{R}^{d \times m}$ is a matrix and $\lambda \in \mathbb{R}_+$ is the regularization parameter. (5 points)

II. Convex Problem

Determine whether these following problems are convex or not respectively.

– If yes, equivalently reformulate the original problem into a standard convex optimization form, i.e., Linear Programming (LP), Second-Order Cone Programming (SOCP) and Semidefinite Programming (SDP).

– If no, relax the original problem to convex problem at first, and then equivalently reformulate the relaxed problem into a standard convex optimization form.

Hint: Standard convex optimization form of Linear Programming (LP):

$$\begin{aligned} \underset{\boldsymbol{x}}{\text{minimize}} \quad & \boldsymbol{c}^\mathsf{T}\boldsymbol{x} + d \\ \text{subject to} \quad & \boldsymbol{G}\boldsymbol{x} \leq \boldsymbol{h} \\ & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}. \end{aligned} \tag{1}$$

Standard convex optimization form of Second-Order Cone Programming (SOCP):

$$\begin{aligned} \underset{\boldsymbol{x}}{\text{minimize}} \quad & \boldsymbol{f}^\mathsf{T}\boldsymbol{x} \\ \text{subject to} \quad & \|\boldsymbol{A}_i\boldsymbol{x} + \boldsymbol{b}_i\| \leq \boldsymbol{c}_i^\mathsf{T}\boldsymbol{x} + d_i, \ i = 1, \dots m \\ & \boldsymbol{F}\boldsymbol{x} = \boldsymbol{g}. \end{aligned} \tag{2}$$

Standard convex optimization form of Semidefinite Programming (SDP):

$$\begin{aligned} \underset{\boldsymbol{x}}{\text{minimize}} \quad & \boldsymbol{c}^\mathsf{T}\boldsymbol{x} \\ \text{subject to} \quad & x_1\boldsymbol{F}_1 + x_2\boldsymbol{F}_2 + \ldots + x_n\boldsymbol{F}_n \preceq \boldsymbol{G} \\ & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}. \end{aligned} \tag{3}$$

1. Consider the following compressive sensing problem

$$\begin{aligned} \underset{\boldsymbol{x}}{\text{minimize}} \quad & \|\boldsymbol{x}\|_1 \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \end{aligned} \tag{4}$$

where $\boldsymbol{A} \in \mathbb{R}^{m \times d}$, $\boldsymbol{x} \in \mathbb{R}^d$ and $\boldsymbol{b} \in \mathbb{R}^m$. (10 points)

2. Consider the following coordinated beamforming design problem

$$\begin{aligned} \underset{\boldsymbol{w}_1, \cdots, \boldsymbol{w}_K}{\text{minimize}} \quad & \sum_{k=1}^{K} \|\boldsymbol{w}_k\|_2^2 \\ \text{subject to} \quad & \frac{\left|\boldsymbol{h}_k^{\mathsf{H}} \boldsymbol{w}_k\right|^2}{\sum_{i \neq k} \left|\boldsymbol{h}_k^{\mathsf{H}} \boldsymbol{w}_i\right|^2 + \sigma^2} \geq \gamma_k, \; k = 1, \ldots, K, \end{aligned} \tag{5}$$

where $\boldsymbol{w}_k \in \mathbb{C}^m$, $\boldsymbol{h}_k \in \mathbb{C}^m$, $\sigma^2 \geq 0$ and $\gamma_k \geq 0$. (10 points)

3. Consider the following problem

$$\begin{aligned} \underset{\boldsymbol{v}}{\text{minimize}} \quad & \boldsymbol{v}^{\top} \boldsymbol{R} \boldsymbol{v} \\ \text{subject to} \quad & |v_i| = 1, \; i = 1, \ldots, m, \end{aligned} \tag{6}$$

where $\boldsymbol{R} \in \mathbb{R}^{m \times m}$ and $\boldsymbol{v} \in \mathbb{R}^m$. (10 points)

III. Water Filling

Consider the following power allocation problem

$$\begin{aligned} \underset{p_1, \ldots, p_K}{\text{maximize}} \quad & \sum_{k=1}^{K} \ln\left(1 + \frac{p_k |h_k|^2}{N_0}\right) \\ \text{subject to} \quad & \sum_{k=1}^{K} p_k = P_{\max} \\ & p_k \geq 0, \; k = 1, \ldots, K, \end{aligned} \tag{7}$$

where $N_0 > 0$.

1. Determine that this problem is convex or not, and provide your argument. (5 points)
2. Write down the dual problem of (7). (5 points)
3. Derive the KKT conditions of (7). (5 points)
4. Derive the expression of the optimal solution of (7). (5 points)

IV. Gradient Descent With Error

Consider unconstrained minimization of the quadratic function $f(\boldsymbol{z}) = \frac{1}{2}\boldsymbol{z}^{\top}\boldsymbol{Q}\boldsymbol{z} - \boldsymbol{b}^{\top}\boldsymbol{z}$ where $\boldsymbol{Q}$ is symmetric and strictly positive definite, and let $\boldsymbol{z}^*$ be the global minimum.

1. Imagine that we can only compute the gradient up to some error, and so we implement the algorithm

$$\boldsymbol{z}^{k+1} = \boldsymbol{z}^k - \alpha\left(\nabla f(\boldsymbol{z}^k) + \boldsymbol{e}^k\right)$$

where $\boldsymbol{e}^k \in \mathbb{R}^d$ is some additive error term, satisfying a bound of the form $\|\boldsymbol{e}^k\|_2 \leq B$. Prove that (with appropriate choice of step size α) the algorithm converges to some ball around $\boldsymbol{z}^*$. Specify the radius of this ball. (10 points)

2. Now suppose that the sequence $\{\boldsymbol{e}^k\}_{k=1}^{\infty}$ consists of independent random vectors, zeros-mean and with expectation $\mathbb{E}\|\boldsymbol{e}^k\|_2^2 \leq S^2$. For a fixed step size $\alpha > 0$, again appropriately chosen, state and prove a bound on the expected mean squared error $\mathbb{E}\|\boldsymbol{z}^k - \boldsymbol{z}^*\|_2^2$. (10 points)
Hint: $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$, when random variable X is independent of random variable Y.

V. Alternating Direction Method of Multipliers

Consider the least squares regression problem with ℓ_2-norm regularization (ridge penalty),

$$\underset{\boldsymbol{z} \in \mathbb{R}^d}{\text{minimize}} \quad \frac{1}{2} \sum_{i=1}^{N} \|\boldsymbol{A}_i \boldsymbol{z} - \boldsymbol{b}_i\|_2^2 + (\lambda/2)\|\boldsymbol{z}\|_2^2 \tag{8}$$

where $\boldsymbol{A}_i \in \mathbb{R}^{n_i \times d}$ is the data matrix, $\boldsymbol{b}_i \in \mathbb{R}^{n_i}$ is the measurement vector, λ is the regularization parameter.

Please write the **exact** ADMM update steps for this problem. (10 points)