SI251 - Convex Optimization, Spring 2021
Homework 2

Due on April 6, 2021, 23:59 UTC+8

1. (*Linear Programming*) Consider the following compressive sensing problem via ℓ_1-minimization:

$$\begin{aligned} \underset{\boldsymbol{x}}{\text{minimize}} \quad & \|\boldsymbol{x}\|_1 \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{z}, \end{aligned} \tag{1}$$

where $\boldsymbol{A} \in \mathbb{R}^{m\times d}$, $\boldsymbol{x} \in \mathbb{R}^d$ and $\boldsymbol{z} \in \mathbb{R}^m$.

(a) Equivalently reformulate (1) into a linear programming problem. (15 points)

Solution:

Suppose unknown signal is component-wise non-negative, ℓ_1 minimization problem is just

$$\begin{aligned} \underset{\boldsymbol{x}\in\mathbb{R}^d}{\text{minimize}} \quad & \sum_{i=1}^{n} \boldsymbol{x}_i \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{z} \\ & \boldsymbol{x} \geq 0. \end{aligned} \tag{2}$$

The general case of real-valued signals, the key trick is to add additional variables to "linearize" the non-linear objective function. Use $\boldsymbol{y}_i$ to represent $\boldsymbol{x}_i$, then we have

$$\begin{aligned} \underset{\boldsymbol{x}\in\mathbb{R}^d}{\text{minimize}} \quad & \sum_{i=1}^{n} \boldsymbol{y}_i \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{z} \\ & \boldsymbol{y}_i = |\boldsymbol{x}_i|, i = 1, 2, \ldots, n. \end{aligned} \tag{3}$$

However, this problem is non-convex due to the second constraints. So we add "linear" inequalities, that is

$$\begin{aligned} \boldsymbol{y}_i - \boldsymbol{x}_i \geq 0, i = 1, 2, \ldots, n \\ \boldsymbol{y}_i + \boldsymbol{x}_i \geq 0, i = 1, 2, \ldots, n, \end{aligned} \tag{4}$$

which is equivalent to

$$\boldsymbol{y}_i \geq \max\{\boldsymbol{x}_i, -\boldsymbol{x}_i\} = |\boldsymbol{x}_i|, i = 1, 2, \ldots, n,$$

then we have the LP problem:

$$\begin{aligned} \underset{\boldsymbol{x}\in\mathbb{R}^d}{\text{minimize}} \quad & \sum_{i=1}^{n} \boldsymbol{y}_i \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{z} \\ & \boldsymbol{y}_i \geq \boldsymbol{x}_i, i = 1, 2, \ldots, n \\ & \boldsymbol{y}_i \geq -\boldsymbol{x}_i, i = 1, 2, \ldots, n. \end{aligned} \tag{5}$$

(b) Write down the dual problem of the reformulated linear program in (a). (15 points)

Solution:

The Lagrangian function of (5) is

$$\begin{aligned} L(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\lambda},\boldsymbol{u},\boldsymbol{v}) &= \sum_{i=1}^{n}\boldsymbol{y}_i + \boldsymbol{\lambda}^T(\boldsymbol{A}\boldsymbol{x}-\boldsymbol{z}) + \boldsymbol{u}^T(\boldsymbol{x}-\boldsymbol{y}) - \boldsymbol{v}^T(\boldsymbol{x}+\boldsymbol{y}) \\ &= (\mathbf{1}-\boldsymbol{u}-\boldsymbol{v})^T\boldsymbol{y} + \left(\boldsymbol{\lambda}^T\boldsymbol{A}+\boldsymbol{u}^T-\boldsymbol{v}^T\right)\boldsymbol{x} - \boldsymbol{\lambda}^T\boldsymbol{z}. \end{aligned} \tag{6}$$

The stationary condition of this function is

$$\begin{aligned} \frac{\partial L}{\partial \boldsymbol{y}} &= \mathbf{1}-\boldsymbol{u}-\boldsymbol{v} = 0, \\ \frac{\partial L}{\partial \boldsymbol{x}} &= \boldsymbol{A}^T\boldsymbol{\lambda}+\boldsymbol{u}-\boldsymbol{v} = 0. \end{aligned} \tag{7}$$

So we have the dual problem of (5):

$$\begin{aligned} \underset{\boldsymbol{\lambda},\boldsymbol{u},\boldsymbol{v}}{\text{maximize}} \quad & -\boldsymbol{\lambda}^T\boldsymbol{z} \\ \text{subject to} \quad & \boldsymbol{u}\geq 0 \\ & \boldsymbol{v}\geq 0. \end{aligned} \tag{8}$$

2. (*Second-Order Cone Programming*) Consider the following coordinated beamforming design problem for transmit power minimization in wireless communication networks [1]

$$\begin{aligned} \mathscr{P}: \underset{\boldsymbol{w}_1,\cdots,\boldsymbol{w}_K}{\text{minimize}} \quad & \sum_{k=1}^{K}\|\boldsymbol{w}_k\|^2 \\ \text{subject to} \quad & \text{SINR}_k(\boldsymbol{w}_1,\cdots,\boldsymbol{w}_K)\geq\gamma_k, k=1,\cdots,K, \end{aligned} \tag{9}$$

where $\boldsymbol{w}_k\in\mathbb{C}^n$ is the transmit beamforming vector for user k, and $\gamma_k\geq 0$ is the threshold for quality-of-service (QoS) requirement. The signal-to-interference-plus-noise-ratio (SINR) for k -th user is given by

$$\text{SINR}_k(\boldsymbol{w}_1,\cdots,\boldsymbol{w}_K) = \frac{\left|\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{w}_k\right|^2}{\sum_{i\neq k}\left|\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{w}_i\right|^2+\sigma^2}, \tag{10}$$

where $\boldsymbol{h}^k\in\mathbb{C}^n$ is the channel coefficient vector between the transmitter and the k -th user and $\sigma^2\geq 0$ is noise power. Parameters $\boldsymbol{h}_k,\gamma_k,\sigma^2$ are known in this problem.

(a) Equivalently reformulate problem $\mathscr{P}$ into a second-order cone programming (SOCP) problem. (20 points)

Solution:

$$\begin{aligned} \underset{\boldsymbol{w}_1,\cdots,\boldsymbol{w}_K}{\text{minimize}} \quad & \sum_{k=1}^{K}\|\boldsymbol{w}_k\|^2 \\ \text{subject to} \quad & \left(1+\frac{1}{\gamma_k}\right)\left|\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{w}_k\right|^2 \geq \sum_{i=1}^{K}\left|\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{w}_i\right|^2+\sigma^2,\ k=1,\cdots,K, \end{aligned} \tag{11}$$

Since $\boldsymbol{w}_k$ and $e^{j\theta_k}\boldsymbol{w}_k$ are completely equivalent for any rotation $\theta\in\mathbb{R}$, we can always assume $\boldsymbol{h}_k^H\boldsymbol{w}_k$ is real and non-negative without loss of generality. Then we can get the following equivalent SOCP reformulation

$$\begin{aligned} \mathscr{P}_{2.1}: \underset{\boldsymbol{w}_1,\cdots,\boldsymbol{w}_K}{\text{minimize}} \quad & \sum_{k=1}^{K}\|\boldsymbol{w}_k\|^2 \\ \text{subject to} \quad & \sqrt{\sum_{i=1}^{K}\left(\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{w}_i\right)^2+\sigma^2} \leq \sqrt{1+\frac{1}{\gamma_k}}\,\text{Real}\{\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{w}_k\},\ k=1,\cdots,K. \end{aligned} \tag{12}$$

(b) Find the global optimal solution to problem $\mathscr{P}$ using Lagrangian duality approach. (20 points)

Solution:

From the SOCP reformulation $\mathscr{P}_{2.1}$, it's easy to show that Slater's condition is fulfilled. Hence, strong duality and KKT conditions are necessary and sufficient for the optimal solution. We define the Lagrangian function as

$$\mathcal{L}(\boldsymbol{w}_1, \cdots, \boldsymbol{w}_K, \lambda_1, \cdots, \lambda_K) = \sum_{j=1}^{K} \|\boldsymbol{w}_j\|_2^2 + \sum_{j=1}^{K} \lambda_j \left(\sum_{i \neq j} \left(\boldsymbol{h}_j^{\mathrm{H}} \boldsymbol{w}_i\right)^2 + \sigma^2 - \frac{1}{\gamma_j} \left(\boldsymbol{h}_j^{\mathrm{H}} \boldsymbol{w}_j\right)^2 \right) \tag{13}$$

From zero gradient of Lagrangian condition, we know $\partial \mathcal{L} / \partial \boldsymbol{w}_k = \boldsymbol{0}$, i.e.

$$2\boldsymbol{w}_k + \sum_{j \neq k} 2\lambda_j \boldsymbol{h}_j \boldsymbol{h}_j^{\mathrm{H}} \boldsymbol{w}_k - 2\lambda_k / \gamma_k \boldsymbol{h}_k \boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{w}_k = \boldsymbol{0} \tag{14}$$

$$\Rightarrow \left(\boldsymbol{I} + \sum_{k=1}^{K} \lambda_k \boldsymbol{h}_k \boldsymbol{h}_k^{\mathrm{H}} \right) \boldsymbol{w}_k = \lambda_k \left(1 + \frac{1}{\gamma_k}\right) \boldsymbol{h}_k \boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{w}_k \tag{15}$$

$$\Rightarrow \boldsymbol{w}_k = \left(\boldsymbol{I} + \sum_{k=1}^{K} \lambda_k \boldsymbol{h}_k \boldsymbol{h}_k^{\mathrm{H}} \right)^{-1} \boldsymbol{h}_k \times \underbrace{\lambda_k \left(1 + \frac{1}{\gamma_k}\right) \boldsymbol{h}_k^{\mathrm{H}} \boldsymbol{w}_k}_{\text{scalar}} \tag{16}$$

The last equation holds since $\boldsymbol{I} + \sum_{k=1}^{K} \lambda_k \boldsymbol{h}_k \boldsymbol{h}_k^{\mathrm{H}} \succ \boldsymbol{0}$ ($\lambda_k \geq 0$ from dual feasibility). Then we know

$$\boldsymbol{w}_k^* = \sqrt{p_k} \frac{\left(\boldsymbol{I} + \sum_{k=1}^{K} \lambda_k \boldsymbol{h}_k \boldsymbol{h}_k^{\mathrm{H}} \right)^{-1} \boldsymbol{h}_k}{\left\| \left(\boldsymbol{I} + \sum_{k=1}^{K} \lambda_k \boldsymbol{h}_k \boldsymbol{h}_k^{\mathrm{H}} \right)^{-1} \boldsymbol{h}_k \right\|_2} \tag{17}$$

where $p_k \geq 0$ and $\sqrt{p_k} = \|\boldsymbol{w}_k\|$. Since (16) always holds for any $p_k \geq 0$, we know the optimal solution of $\mathscr{P}$ is given by

$$\boldsymbol{w}_k^* = \underbrace{\sqrt{p_k}}_{\text{beamforming power}} \underbrace{\frac{\left| \left(\boldsymbol{I} + \sum_{k=1}^{K} \lambda_k \boldsymbol{h}_k \boldsymbol{h}_k^{\mathrm{H}} \right)^{-1} \boldsymbol{h}_k \right|}{\left\| \left(\boldsymbol{I} + \sum_{k=1}^{K} \lambda_k \boldsymbol{h}_k \boldsymbol{h}_k^{\mathrm{H}} \right)^{-1} \boldsymbol{h}_k \right\|_2}}_{\text{beamforming direction}} \underbrace{e^{j\theta_k}}_{\text{phase ambiguity}}, k = 1, \cdots, K \tag{18}$$

where $\sqrt{p_k} = \|\boldsymbol{w}_k\|$, and $\theta_k \in \mathbb{R}$ is the phase of $\boldsymbol{w}_k^*$.

3. (*Semidefinite Programming*) Consider $\boldsymbol{A}(\boldsymbol{x}) = \boldsymbol{A}_0 + x_1 \boldsymbol{A}_1 + \cdots + x_n \boldsymbol{A}_n$, where the vector $\boldsymbol{x} \in \mathbb{R}^n$ and the matrix $\boldsymbol{A}_i \in \mathbb{S}^m$, for $i = 0, 1, \cdots, n$. Let $\lambda_1(\boldsymbol{x}) \geq \cdots \geq \lambda_m(\boldsymbol{x})$ denotes the eigenvalues of $\boldsymbol{A}(\boldsymbol{x})$. Equivalently reformulate the following problems as SDPs.

(a) $\min_{\boldsymbol{x}} \quad \lambda_1(\boldsymbol{x})$. (5 points)

Solution:

$\lambda_1(\boldsymbol{x}) \leq t$ if and only if $\boldsymbol{A}(\boldsymbol{x}) \preceq t\boldsymbol{I}$, so we have

$$\begin{aligned} \underset{\boldsymbol{x}, t}{\text{minimize}} \quad & t \\ \text{subject to} \quad & \boldsymbol{A}(\boldsymbol{x}) \preceq t\boldsymbol{I}. \end{aligned} \tag{19}$$

(b) $\min_{\boldsymbol{x}} \quad \lambda_1(\boldsymbol{x}) - \lambda_m(\boldsymbol{x})$. (10 points)

Solution:

$\lambda_1(\boldsymbol{x}) \leq t_1$ if and only if $\boldsymbol{A}(\boldsymbol{x}) \preceq t_1 \boldsymbol{I}$ and $\lambda_m(\boldsymbol{x}) \geq t_2$ if and only if $\boldsymbol{A}(\boldsymbol{x}) \succeq t_2 \boldsymbol{I}$, so we have

$$\begin{aligned} \underset{\boldsymbol{x}, t_1, t_2}{\text{minimize}} \quad & t_1 - t_2 \\ \text{subject to} \quad & t_2 \boldsymbol{I} \preceq \boldsymbol{A}(\boldsymbol{x}) \preceq t_1 \boldsymbol{I}. \end{aligned} \tag{20}$$

(c) $\min_{\boldsymbol{x}} \quad \sum_{i=1}^{m} |\lambda_i(\boldsymbol{x})|$. (15 points)

Solution:

Method 1:

Suppose $\boldsymbol{A}(\boldsymbol{x})$ has eigenvalue decompositon $\boldsymbol{Q\Lambda Q}^T$. Let $\boldsymbol{A}(\boldsymbol{x}) = \boldsymbol{A}^+ - \boldsymbol{A}^- = \boldsymbol{Q\Lambda}^+\boldsymbol{Q}^T - \boldsymbol{Q\Lambda}^-\boldsymbol{Q}^T$.$\boldsymbol{\Lambda}$ is divided into two parts: $\boldsymbol{\Lambda}^+$ and $\boldsymbol{\Lambda}^- \cdot \lambda_i(\boldsymbol{x}) \geq 0$ are in the $\boldsymbol{\Lambda}^+$ and $-\lambda_i(\boldsymbol{x}) \geq 0$ are in the $\boldsymbol{\Lambda}^-$. Thus $\boldsymbol{A}^+ \succeq 0$ and $\boldsymbol{A}^- \succeq 0$. $\sum_{i=1}^{m} |\lambda_i(\boldsymbol{x})|$ is equivalent to $\operatorname{trace}(\boldsymbol{A}^+) + \operatorname{trace}(\boldsymbol{A}^-)$.

$$\begin{aligned} \underset{\boldsymbol{x},\boldsymbol{A}^+,\boldsymbol{A}^-}{\text{minimize}} \quad & \operatorname{trace}\left(\boldsymbol{A}^+\right) + \operatorname{trace}\left(\boldsymbol{A}^-\right) \\ \text{subject to} \quad & \boldsymbol{A}(\boldsymbol{x}) = \boldsymbol{A}^+ - \boldsymbol{A}^- \\ & \boldsymbol{A}^+ \succeq 0 \\ & \boldsymbol{A}^- \succeq 0 \end{aligned} \tag{21}$$

Method 2:

Similarly to ℓ_1 norm of vector, we have

$$\begin{aligned} \underset{\boldsymbol{x},\boldsymbol{Y}}{\text{minimize}} \quad & \operatorname{trace}(\boldsymbol{Y}) \\ \text{subject to} \quad & \boldsymbol{Y} + \boldsymbol{A}(\boldsymbol{x}) \succeq 0 \\ & \boldsymbol{Y} - \boldsymbol{A}(\boldsymbol{x}) \succeq 0. \end{aligned} \tag{22}$$

References

[1] E. Björnson, M. Bengtsson, and B. Ottersten, "Optimal multiuser transmit beamforming: A difficult problem with a simple solution structure [lecture notes]," *IEEE Signal Process. Mag.*, vol. 31, pp. 142–148, Jul 2014.