MAT3220: Operation Research
Homework 8

李肖鹏 (116010114)

Due date: April 30, 12 p.m., 2019

Problem 1. Show that the following compressive sensing model

$$\begin{aligned} \min_{\vec{x}} \quad & \|\vec{x}\|_1 \\ s.t. \quad & A\vec{x} = \vec{b} \end{aligned}$$

can be modeled as a linear programming problem.

Since $\|\vec{x}\|_1 = \sum_{i=1}^{n} |x_i|$, let $|x_i| = t_i$ for all $i = 1, \ldots, n$, then we have

$$\begin{aligned} \min_{x_i, t_i} \quad & \sum_{i=1}^{n} t_i \\ s.t. \quad & A\vec{x} = \vec{b} \\ & |x_i| = t_i, \quad \forall\, i = 1, \ldots, n \end{aligned}$$

which can be further relaxed into

$$\begin{aligned} \min_{x_i, t_i} \quad & \sum_{i=1}^{n} t_i \\ s.t. \quad & A\vec{x} = \vec{b} \\ & |x_i| \le t_i, \quad \forall\, i = 1, \ldots, n \end{aligned}$$

Let $\vec{t} = (t_1, \cdots, t_n)^{\mathrm{T}}$ and $\vec{e} = (1, \ldots, 1)^{\mathrm{T}}$, we have

$$\begin{aligned} \min_{\vec{x}, \vec{t}} \quad & \vec{e}^{\mathrm{T}} \vec{t} \\ s.t. \quad & A\vec{x} = \vec{b} \\ & \vec{x} \le \vec{t} \\ & \vec{t} \ge -\vec{x} \end{aligned}$$

Problem 2.

- Derive the KKT optimality condition for the following SVM (Support Vector Machine) model (you don't need to solve it),

$$\begin{aligned} \min_{\vec{w},b} \quad & \frac{1}{2}\|\vec{w}\|^2 \\ s.t. \quad & \vec{w}^{\mathrm{T}}\vec{x}_i - b \geq 1, \; i \in I \\ & \vec{w}^{\mathrm{T}}\vec{x}_j - b \leq 1, \; j \in J \end{aligned}$$

First, we write down the Lagrangian function,

$$L(\lambda_i, \lambda_j, \vec{w}, b) = \frac{1}{2}\|\vec{w}\|^2 + \sum_{i \in I} \lambda_i(-\vec{x}_i^{\mathrm{T}}\vec{w} + b - 1) + \sum_{j \in J} \lambda_j(\vec{x}_j^{\mathrm{T}}\vec{w} - b - 1)$$

The KKT condition says that there exists $\vec{w}^*$, b^*, λ_i^* and λ_j^* (for $i \in I, j \in J$) such that

$$\vec{w}^* + \sum_{j \in J} \lambda_j^* \vec{x}_j = \sum_{i \in I} \lambda_i^* \vec{x}_i, \qquad \sum_{i \in I} \lambda_i^* = \sum_{j \in J} \lambda_j^*$$

$$\lambda_i^*(-\vec{x}_i^{\mathrm{T}}\vec{w}^* + b^* - 1) = 0, \quad \lambda_j^*(\vec{x}_j^{\mathrm{T}}\vec{w}^* - b^* - 1) = 0, \quad \forall i \in I, j \in J$$

$$\lambda_i^* \geq 0, \quad \lambda_j^* \geq 0, \quad \forall i \in I, j \in J$$

$$\vec{x}_i^{\mathrm{T}}\vec{w}^* - b^* \geq 1, \; \forall i \in I, \qquad \vec{x}_j^{\mathrm{T}}\vec{w}^* - b^* \leq 1, \; \forall j \in J$$

- Prove that the logistic regression function

$$F(\vec{w}, b) = \sum_{i=1}^{m} \ln\left(1 + \exp(-s_i(\vec{w}^{\mathrm{T}}\vec{x}_i - b))\right)$$

is convex, where $\vec{x}_i$'s are the given data vectors to learn, and $s_i \in \{-1, +1\}$ are the given identifiers. (In other words, the variables are $(\vec{w}, b)$.)

To prove F is convex with respect to $(\vec{w}, b)$, it suffices to show that for any $i = 1, \ldots, m$,

$$f(\vec{w}, b) = \ln\left(1 + \exp(-s_i(\vec{w}^{\mathrm{T}}\vec{x}_i - b))\right)$$

is convex in $(\vec{w}, b)$. Let $g(x) : \mathbb{R} \mapsto \mathbb{R}$ defined by $g(x) = \ln(1 + e^x)$, then it is easy to show that $g(x)$ is convex, since $g'(x) = \frac{e^x}{1+e^x}$, and $g''(x) = \frac{e^x}{(1+e^x)^2} > 0$.

Next, let $h_i(\vec{w}, b) = -s_i(\vec{w}^{\mathrm{T}}\vec{x}_i - b)$, we claim that $h(\vec{w}, b)$ is linear function, i.e., for any $p, q \in \mathbb{R}$, $\vec{w}, \vec{w}'$, and b, b', we have

$$\begin{aligned} h(p(\vec{w}, b) + q(\vec{w}', b')) &= h(p\vec{w} + q\vec{w}', pb + qb') \\ &= -s_i((p\vec{w} + q\vec{w}')^{\mathrm{T}}\vec{x}_i - (pb + qb')) \\ &= p[-s_i(\vec{w}^{\mathrm{T}}\vec{x}_i - b)] + q[-s_i(\vec{w}'^{\mathrm{T}}\vec{x}_i - b')] \\ &= ph(\vec{w}, b) + qh(\vec{w}', b') \end{aligned}$$

Finally, we prove that the composite function of convex function and linear function is still convex. Let $\phi : \mathbb{R} \mapsto \mathbb{R}$ be convex, and $\psi : \mathbb{R}^n \mapsto \mathbb{R}$ is linear, then

$$\begin{aligned} \phi(\psi(\lambda\vec{x} + (1-\lambda)\vec{y})) &= \phi(\lambda\psi(\vec{x}) + (1-\lambda)\psi(\vec{y})) \\ &\leq \lambda\phi(\psi(\vec{x})) + (1-\lambda)\phi(\psi(\vec{y})) \end{aligned}$$

which means the composite of them is convex, therefore, $f(\overrightarrow{w}, b) = g(h_i(\overrightarrow{w}, b))$ is convex, and this further implies that $F(\overrightarrow{w}, b)$ is convex because the sum of convex function is also convex.