Advanced Topics in Control 2020: Large-Scale Convex Optimization

Exercise 3: Convex Optimization Problems

Goran Banjac, Mathias Hudoba de Badyn, Andrea Iannelli,
Angeliki Kamoutsi, Ilnura Usmanova

March 27, 2020

Date due: March 26, 2019 at 23:59.

Please submit your solutions via Moodle as a PDF with filename `Ex01_Surname.pdf`, replacing `Surname` with your surname.

---

1 Local Minima Are Global Minima in Convex Optimization

Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ is a convex function and assume that $x^\star$ is a local minimizer of f, i.e., there exists a sufficiently small $\delta > 0$ such that for all x with $\|x - x^\star\| \leq \delta$, it holds that $f(x^\star) \leq f(x)$. Show that $x^\star$ is a global minimizer of f.
Hint: Assume that $x^\star$ is not a global minimizer., i.e., there exists $\overline{x} \in \mathbb{R}^n$, such that $f(\overline{x}) < f(x^\star)$. In order to arrive to a contradiction, consider the point $x = \theta\overline{x} + (1-\theta)x^\star$ for sufficiently small θ so that $\|x - x^\star\| \leq \delta$ (for instance $\theta = \min\{1, \frac{\delta}{\|x^* - \overline{x}\|}\}$).

2 The Lasso Problem

Consider the *Lasso problem*

$$\min_{x \in \mathbb{R}^n} \frac{1}{2}\|Ax - b\|_2^2 + \lambda\|x\|_1,$$

where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ and $\lambda > 0$. Show that if $\lambda \geq \|A^\top b\|_\infty$, then $x = 0$ is an optimal solution (minimizer).
Hint: Set $f(x) = \frac{1}{2}\|Ax - b\|_2^2 + \lambda\|x\|_1$. Clearly, $f(x) \geq \frac{1}{2}\|Ax - b\|_2^2 + \|A^\top b\|_\infty\|x\|_1$, for all $x \in \mathbb{R}^n$. Continue the argument to derive the lower bound $f(x) \geq \frac{1}{2}\|b\|_2^2$ and show that it is attained at $x = 0$. In the lower bound derivation you can use the inequality $|x^\top y| \leq \|x\|_\infty\|y\|_1$, for all $x, y \in \mathbb{R}^n$.

3 Logistic Regression

(a) Given some data points $x_i \in \mathbb{R}^n$ of some class $y_i \in \{-1, 1\}$ we model the probability of some data point x belonging to class $y = 1$ or $y = -1$ with the following logistic model:

$$\begin{aligned}
\text{Prob}(y = 1) &= p_1(x) = \frac{1}{1 + e^{-(w^\top x + b)}}, \\
\text{Prob}(y = -1) &= p_{-1}(x) = \frac{1}{1 + e^{(w^\top x + b)}},
\end{aligned}$$

where $w \in \mathbb{R}^n$ and $b \in \mathbb{R}$ are the model parameters. With this model, the likelihood for measuring the data (x_i, y_i) for $i = 1, \ldots, N$ is

$$l_w(x,y) = \prod_{i=1}^{N} p_{y_i}(x_i).$$

The model parameters w and b should be chosen such that this likelihood is maximized.

Show that the *maximum likelihood estimate* of (w, b) is given by the following *logistic regression* problem

$$\min_{w,b} \sum_{i=1}^{N} \log(1 + e^{-y_i(x_i^\top w + b)}). \tag{1}$$

Is this problem convex? μ-strongly convex?
Hint: First note that since $y \in \{-1, 1\}$, we can write $p_y(x) = \frac{1}{1+e^{-y(w^\top x+b)}}$. Maximizing $l_w(x,y)$ is the same as maximizing $\log(l_w(x,y))$ (why?). The final optimization problem is not strongly convex. Argue on this by considering the second derivative of $\log(1+e^x)$ as $x \to -\infty$ and $x \to +\infty$.

(b) Show that the problem is equivalent to

$$\min_{w,b} \sum_{i=1}^{N} \left(\log(1 + e^{x_i^\top w + b}) - y_i(x_i^\top w + b) \right) \tag{2}$$

if the classes are labeled with $\{0, 1\}$ instead of $\{-1, 1\}$.
Hint: We have

$$\begin{aligned} \sum_{i=1}^{N} \log(1 + e^{-y_i(w^\top x_i + b)}) &= \sum_{i:y_i=-1} \log(1 + e^{w^\top x_i + b}) + \sum_{i:y_i=1} \log(1 + e^{-(w^\top x_i + b)}) \\ &= \ldots \\ &= \sum_{i=1}^{N} \log(1 + e^{w^\top x_i + b}) - \sum_{i:y_i=1} (w^\top x_i + b). \end{aligned}$$

In the above calculations use that $\log(1 + e^{-(w^\top x_i + b)}) = \log(\frac{1+e^{w^\top x_i+b}}{e^{w^\top x_i+b}})$. Next, go over the new labels.

(c) Consider the logistic regression problem (1) where $y_i \in \{-1, 1\}$ are labels. Assume that there exists $(\overline{w}, \overline{b})$ such that $x_i^\top \overline{w} + \overline{b} < 0$, for all i with $y_i = -1$ and $x_i^\top \overline{w} + \overline{b} > 0$, for all i with $y_1 = 1$. Show that optimal value is 0 and that no (w, b) exists that attains the value.
Hint: First note that the objective function

$$\sum_{i=1}^{N} \underbrace{\log(1 + e^{-y_i(x_i^\top w + b)})}_{=:f_i(w,b)}$$

is striclty positive everywhere, since this is the case for all terms f_i. We want to show that the infimum is 0. To this end, find a sequence $((w_n, b_n))_n$ so that $\lim_{n\to\infty} f_i(w_n, b_n) = 0$, for all $i = 1, \ldots, N$. Treat seperately the i's such that $y_i = -1$ and those that $y_i = 1$.

4 ℓ_1-, ℓ_∞- and ℓ_4-Norm Approximation Problems

(a) Formulate the following problems as LPs. Explain in detail the relation between the optimal solution of each problem and the solution of its equivalent LP.

(1) Minimize $\|Ax - b\|_\infty$.
Hint: Recall the epigraph reformulation discussed on Slide 6 of Lecture 4.

(2) Minimize $\|Ax - b\|_1$.
Hint: The objective (cost function) is $\mathbf{1}^\top s$, where $s \in \mathbb{R}^m$ is a new (additional) optimization variable and $\mathbf{1}$ is the vector with all entries equal to one.

(3) Minimize $\|Ax - b\|_1$ subject to $\|x\|_\infty \leq 1$.

(4) Minimize $\|x\|_1$ subject to $\|Ax - b\|_\infty \leq 1$.

(5) Minimize $\|Ax - b\|_1 + \|x\|_\infty$.

In each problem $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$ are given.

(b) (**Bonus Question**) Formulate the following problem as a QCQP: Minimize $\|Ax - b\|_4$.