CSCI 6968 Weekly Participation 3

Shiuli Subhra Ghosh

February 3, 2023

1 Question

All the optimization problems we have seen for fitting ML models so far are convex in nature. We need to follow the general template for proving the OLS problem is convex.
The mathematical formulation of OLS problem is,

$$\min_{\mathbb{R}^d} \frac{1}{n}\|\mathbf{X}\beta - \mathbf{y}\|_2^2 \tag{1}$$

By definition f is convex on its domain if for any $x, y \in \mathrm{dom}(f)$ and any $\alpha \in [0, 1]$ it it the case that,

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y) \tag{2}$$

1. Nonnegative multiples of convex functions are convex.
Solution:
Let, $f(x)$ is a convex function. So, by definition, $f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y)$. Then we need to show, $af(x)$ is also convex, where $a \in \mathbb{R} \cup \{0\}$.

$$\begin{aligned} af(\alpha x + (1-\alpha)y) &\leq a(\alpha f(x) + (1-\alpha)f(y)) \\ &\leq \alpha a f(x) + (1-\alpha)af(y)) \end{aligned} \tag{3}$$

2. Affine functions of the form $f(x) = \langle a, x\rangle + b$ are convex.
Solution:

$$\begin{aligned} f(\alpha x + (1-\alpha)y) &= \langle a, \alpha x + (1-\alpha)y + b\rangle \\ &= \langle a, \alpha x\rangle + \langle a, (1-\alpha)y\rangle + \alpha b + (1-\alpha)b \\ &= \alpha(\langle a, x\rangle + b) + (1-\alpha)(\langle a, y\rangle + b) \\ &= \alpha f(x) + (1-\alpha)f(y) \end{aligned} \tag{4}$$

3. If g is convex and f is affine, then $g(f(.))$ is also convex.
Solution:
From 4 we know that the affine function f is convex. So,

$$\begin{aligned} g(f(\alpha x + (1-\alpha)y)) &= g(\alpha f(x) + (1-\alpha)f(y)) \\ &\leq \alpha g(f(x)) + (1-\alpha)g(f(y)) \end{aligned} \tag{5}$$

Also, domain of $g(f(x))$ is $f(x)$ which is convex. So, $g(f(x))$ is convex.

4. Sum of convex functions is also convex.
Solution:
Let's take n convex functions $f_i(x), i = 1 : n, n \in \mathbb{N}$. All the f_i s are convex.
So, $f_i(\alpha x + (1-\alpha)y) \leq \alpha f_i(x) + (1-\alpha) f_i(y)$

$$\sum_{i=1}^{n} f_i(\alpha x + (1-\alpha)y) \leq \alpha \sum_{i=1}^{n} f_i(x) + (1-\alpha) \sum_{i=1}^{n} f_i(y) \tag{6}$$

Hence, coming to the final problem. OLS,

$$\min_{\mathbb{R}^d} \frac{1}{n} \|\mathbf{X}\beta - \mathbf{y}\|_2^2 \tag{7}$$

1. defined on $\mathbb{R}^d$ which is a convex set. So, the domain is the OLS is convex.

2. $(\mathbf{X}\beta - \mathbf{y})$ is affine function. From 4 it is also convex.

3. We have proved in class that $\|.\|_2$ (2 - norm) is convex.

4. From 5, the composition of convex and affine functions are convex. So, $\|\mathbf{X}\beta - \mathbf{y}\|_2$ is convex. It is the composition of 2 - norm and affine functions.

5. Next, x^2 is convex. and again by 5, we can tell that square of the composition of 2 - norm and affine function are also convex.

6. Finally, $\frac{1}{n}$ is nonnegative. So, from 1 we can say that $\frac{1}{n}\|\mathbf{X}\beta - \mathbf{y}\|_2^2$ is convex.

Hence, OLS is a convex optimization problem.