EPFL CS439 Exercises

I am not affiliated with EPFL, I am just solving the exercises for self-study. Course material: EPFL CS439 GitHub

Siddhartha Bhattacharya

Week 1

Problem 1

(Exercise 2) Prove Jensen's inequality.

Lemma 0.1 (Jensen's Inequality)**.** *Let f be convex, $x_1, \ldots, x_m \in \mathbf{dom}(f)$, $\lambda_1, \ldots, \lambda_m \in \mathbb{R}_{\geq 0}$ such that $\sum_{i=1}^m \lambda_i = 1$. Then*

$$f\left(\sum_{i=1}^{m} \lambda_i x_i\right) \leq \sum_{i=1}^{m} \lambda_i f(x_i) \tag{1}$$

Proof. First, since f is convex, we have that for any $x, y \in \mathbf{dom}(f)$ and $0 \leq \lambda \leq 1$,

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)$$

We proceed by induction. For the base case, let $k = 2$ and $0 \leq \lambda_1, \lambda_2 \leq 1$ such that $\lambda_1 + \lambda_2 = 1$. Note that this implies $\lambda_2 = 1 - \lambda_1$.

$$\begin{aligned} f\left(\sum_{i=1}^{2} \lambda_i x_i\right) =& f(\lambda_1 x_1 + \lambda_2 x_2) \\ =& f(\lambda_1 x_1 + (1-\lambda_1) x_2) \end{aligned}$$

Next, since f is convex,

$$f(\lambda_1 x_1 + (1-\lambda_1) x_2) \leq \lambda_1 f(x_1) + (1-\lambda_1) f(x_2)$$

Now, substituting $\lambda_2 = (1 - \lambda_1)$, we get that

$$f(\lambda_1 x_1 + \lambda_2 x_2) \leq \lambda_1 f(x_1) + \lambda_2 f(x_2)$$

So, (8) holds for $k = 2$. Now, suppose for induction that (8) holds for an arbitrary $k \geq 2$. We show that it holds for $k + 1$. Let $1 = \sum_{i=1}^{k+1} \lambda_i$. If we let $\beta = \sum_{i=1}^{k} \lambda_i$, then $\lambda_{k+1} = 1 - \beta$. Using this, we can rewrite the convex combination $\sum_{i=1}^{k+1} \lambda_i x_i$ as:

$$\begin{aligned} \sum_{i=1}^{k+1} \lambda_i x_i =& \beta \left(\sum_{i=1}^{k} \frac{\lambda_i}{\beta} x_i\right) + \lambda_{k+1} x_{k+1} \\ =& \beta \left(\sum_{i=1}^{k} \frac{\lambda_i}{\beta} x_i\right) + (1-\beta) x_{k+1} \end{aligned}$$

And since f is convex,

$$f\left(\beta \left(\sum_{i=1}^{k} \frac{\lambda_i}{\beta} x_i\right) + (1-\beta) x_{k+1}\right) \leq \beta f\left(\sum_{i=1}^{k} \frac{\lambda_i}{\beta} x_i\right) + (1-\beta) f(x_{k+1})$$

And by inductive hypothesis on the first k terms, we know that $f\left(\sum_{i=1}^{k} \frac{\lambda_i}{\beta} x_i\right) \leq \sum_{i=1}^{k} \frac{\lambda_i}{\beta} f(x_i)$, which is a valid convex combination, since we defined $\beta = \sum_{i=1}^{k} \lambda_i$, implying that $\sum_{i=1}^{k} \frac{\lambda_i}{\beta} = \frac{\sum_{i=1}^{k} \lambda_i}{\beta} = \frac{\beta}{\beta} = 1$.

Problem 1 continued on next page...

Thus,

$$\beta f\left(\sum_{i=1}^{k}\frac{\lambda_i}{\beta}x_i\right)+(1-\beta)f(x_{k+1})\leq\beta\left(\sum_{i=1}^{k}\frac{\lambda_i}{\beta}f(x_i)\right)+(1-\beta)f(x_{k+1})$$

And by the definition of the convex combination, $\beta\left(\sum_{i=1}^{k}\frac{\lambda_i}{\beta}f(x_i)\right)+(1-\beta)f(x_{k+1})=\sum_{i=1}^{k+1}\lambda_i f(x_i)$, we are done. □

Problem 2

(Exercise 4) Prove that the function $d_y:\mathbb{R}^d\to\mathbb{R},\ x\mapsto\|x-y\|^2$ is strictly convex for any $y\in\mathbb{R}^d$ (use Lemma 1.25)

Proof. First, we show that d_y is twice continually differentiable. Let $x\in\mathbb{R}^d$. Then,

$$\begin{aligned}d_y(x)=&\|x-y\|^2=(x-y)^T(x-y)=x^Tx-2x^Ty+y^Ty\\ \nabla d_y(x)=&2x-2y\\ \nabla^2 d_y(x)=&2I\end{aligned}$$

Thus, d_y is twice continually differentiable. Furthermore, $z^T\nabla^2 f(x)z=2z^Tz>0$ for all $z\neq 0$, so $\nabla^2 f(x)$ is positive definite. Since $f(x)$ is twice continually differentiable and $\nabla^2 f(x)$ is positive definite, $f(x)$ is strictly convex, by Lemma 1.25. □