Homework 3

20213073 Donggyu Kim

April 30, 2021

1. (a) True
   (b) True
   (c) True
   (d) True
   (e) True

2. For each following function, $\mathrm{dom}(f) = \mathbb{R}^n$ is a convex set. Therefore, we only need to show the Jensen inequality holds to prove that each function is convex.

   (a) $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n,\ \forall t \in [0, 1]$,

$$\begin{aligned} f(t\boldsymbol{x} + (1-t)\boldsymbol{y}) &= \max(tx_1 + (1-t)y_1, \dots, tx_n + (1-t)y_n) \\ &\leq t \max(x_1, \dots, x_n) + (1-t) \max(y_1, \dots, y_n) \\ &= tf(\boldsymbol{x}) + (1-t)f(\boldsymbol{y}) \end{aligned}$$

   (b) $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n,\ \forall t \in [0, 1]$,

$$\begin{aligned} f(t\boldsymbol{x} + (1-t)\boldsymbol{y}) &= \|t\boldsymbol{x} + (1-t)\boldsymbol{y}\|_p \\ &\leq \|t\boldsymbol{x}\|_p + \|(1-t)\boldsymbol{y}\|_p && (\because \text{Subadditivity}) \\ &= t\|\boldsymbol{x}\|_p + (1-t)\|\boldsymbol{y}\|_p && (\because \text{Homogeneity}) \\ &= tf(\boldsymbol{x}) + (1-t)f(\boldsymbol{y}) \end{aligned}$$

   (c) $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n,\ \forall t \in [0, 1]$,

$$\begin{aligned} f(t\boldsymbol{x} + (1-t)\boldsymbol{y}) &= \log\left(\sum_{i=1}^n e^{tx_i + (1-t)y_i}\right) \\ &= \log\left(\sum_{i=1}^n e^{tx_i} e^{(1-t)y_i}\right) \\ &\leq \log\left[\left\{\sum_{i=1}^n \left(e^{tx_i}\right)^{\frac{1}{t}}\right\}^t \left\{\sum_{i=1}^n \left(e^{(1-t)y_i}\right)^{\frac{1}{1-t}}\right\}^{1-t}\right] && (\because \text{Hölder's inequality}) \\ &= t\log\left(\sum_{i=1}^n e^{x_i}\right) + (1-t)\log\left(\sum_{i=1}^n e^{y_i}\right) \\ &= tf(\boldsymbol{x}) + (1-t)f(\boldsymbol{y}) \end{aligned}$$

3. (a) *Proof.*
$\Rightarrow$) $\forall x, y \in \mathrm{dom}(f),\ \forall t \in (0, 1]$,

$$\begin{aligned} f(y) &= \frac{tf(y) + (1-t)f(x) - (1-t)f(x)}{t} \\ &\geq \frac{f(ty + (1-t)x) - (1-t)f(x)}{t} && (\because \text{Convexity}) \\ &= f(x) + \frac{f(x + t(y-x)) - f(x)}{t} \end{aligned}$$

Since f is differentiable, we get

$$\begin{aligned} f(y) &= \lim_{t \to 0+} f(y) \\ &\geq \lim_{t \to 0+} \left(f(x) + \frac{f(x + t(y-x)) - f(x)}{t} \right) \\ &= f(x) + f'(x)(y - x) \end{aligned}$$

$\Leftarrow$) Let $z = tx + (1-t)y$ for any $x, y \in \mathrm{dom}(f),\ t \in [0, 1]$. By the inequality,

$$\begin{aligned} tf(x) + (1-t)f(y) &\geq t(f(z) + f'(z)(x - z)) + (1-t)(f(z) + f'(z)(y - z)) \\ &= f(z) + f'(z)(tx + (1-t)y - z) \\ &= f(z) = f(tx + (1-t)y) \end{aligned}$$

which means that f is convex. $\square$

(b) *Proof.*
As we already proved (a), let's prove that $\forall x, y \in \mathrm{dom}(f),\ f(y) \geq f(x) + f'(x)(y-x) \iff \forall z \in \mathrm{dom}(f),\ f''(z) \geq 0$ to solve this problem.
$\Rightarrow$) $\forall x, y \in \mathrm{dom}(f)$,

$$\begin{aligned} f'(x)(y - x) &= f(x) + f'(x)(y - x) - f(x) \\ &\leq f(y) - f(x) \\ &\leq f(y) - (f(y) + f'(y)(x - y)) \\ &= f'(y)(y - x) \end{aligned}$$

Thus, we have

$$\begin{aligned} f''(x) &= \lim_{y \to x} \frac{f'(y) - f'(x)}{y - x} \\ &= \lim_{y \to x} \frac{(f'(y) - f'(x))(y - x)}{(y - x)^2} \\ &\geq 0 \end{aligned}$$

$\Leftarrow$) By the mean value version of the Taylor's theorem,

$$\forall x, y \in \mathrm{dom}(f),\ \exists z \in [x, y] \quad \text{s.t.} \quad f(y) = f(x) + f'(x)(y - x) + \frac{1}{2} f''(z)(y - x)^2$$

Since $f''(z) \geq 0$, we can get

$$f(y) \geq f(x) + f'(x)(y - x)$$

Thus, f is convex iff $\forall z \in \mathrm{dom}(f),\ f''(z) \geq 0$. $\square$

4. The Lagrangian of the primal problem is as follows.

$$\mathcal{L}(\boldsymbol{w}, b, \boldsymbol{\lambda}) = \frac{1}{2}\boldsymbol{w}^\top \boldsymbol{w} + \sum_{i=1}^{N} \lambda_i \{1 - y_i(\boldsymbol{w}^\top \boldsymbol{x}_i + b)\}$$

Then we can rewrite the primal problem with the Lagrangian.

$$\begin{aligned} \underset{\boldsymbol{w}, b}{\text{minimize}} \quad & \sup_{\boldsymbol{\lambda}} \mathcal{L}(\boldsymbol{w}, b, \boldsymbol{\lambda}) \\ \text{s.t.} \quad & \boldsymbol{\lambda} \geq \mathbf{0} \end{aligned}$$

We can use $\frac{\partial \mathcal{L}}{\partial \boldsymbol{w}} = 0$ and $\frac{\partial \mathcal{L}}{\partial b} = 0$ to derive $\inf_{\boldsymbol{w}, b} \mathcal{L}(\boldsymbol{w}, b, \boldsymbol{\lambda})$.

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \boldsymbol{w}} &= \boldsymbol{w}^\top - \sum_{i=1}^{N} \lambda_i y_i \boldsymbol{x}_i^\top \Rightarrow \boldsymbol{w} = \sum_{i=1}^{N} \lambda_i y_i \boldsymbol{x}_i \\ \frac{\partial \mathcal{L}}{\partial b} &= -\sum_{i=1}^{N} \lambda_i y_i = 0 \Rightarrow \sum_{i=1}^{N} \lambda_i y_i = 0 \end{aligned}$$

By using these conditions, we have

$$\begin{aligned} \inf_{\boldsymbol{w}, b} \mathcal{L}(\boldsymbol{w}, b, \boldsymbol{\lambda}) &= \frac{1}{2}\boldsymbol{w}^\top \boldsymbol{w} + \sum_{i=1}^{N} \lambda_i - \boldsymbol{w}^\top \sum_{i=1}^{N} \lambda_i y_i \boldsymbol{x}_i - \left(\sum_{i=1}^{n} \lambda_i y_i\right) b \\ &= \frac{1}{2}\boldsymbol{w}^\top \boldsymbol{w} + \sum_{i=1}^{N} \lambda_i - \boldsymbol{w}^\top \boldsymbol{w} \\ &= -\frac{1}{2}\left(\sum_{i=1}^{N} \lambda_i y_i \boldsymbol{x}_i\right)^\top \left(\sum_{i=1}^{N} \lambda_i y_i \boldsymbol{x}_i\right) + \sum_{i=1}^{N} \lambda_i \\ &= -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} \lambda_i \lambda_j y_i y_j \boldsymbol{x}_i^\top \boldsymbol{x}_j + \sum_{i=1}^{N} \lambda_i \end{aligned}$$

Thus, the Lagrangian dual problem of the primal is as below.

$$\begin{aligned} \underset{\boldsymbol{\lambda}}{\text{maximize}} \quad & \sum_{i=1}^{N} \lambda_i - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} \lambda_i \lambda_j y_i y_j \boldsymbol{x}_i^\top \boldsymbol{x}_j \\ \text{s.t.} \quad & \boldsymbol{\lambda} \geq \mathbf{0}, \\ & \sum_{i=1}^{N} \lambda_i y_i = 0 \end{aligned}$$

There exist $\boldsymbol{w}$ and b that satisfy $1 - y_i(\boldsymbol{w}^\top \boldsymbol{x}_i + b) \leq 0$ for $i = 1, \ldots, N$. Thus, $(2\boldsymbol{w}, 2b)$ is strictly feasible; $1 - y_i((2\boldsymbol{w})^\top \boldsymbol{x}_i + (2b)) \leq -1 < 0$. Since the convex problem satisfies the Slater's condition, the strong duality holds.