Homework 7 Solutions :: MATH 504

Your homework submission must be a single pdf called "LASTNAME-hw5.pdf" with your solutions to all theory problem to receive full credit. All answers must be typed in Latex.

1. Consider the sets

$$C = \{(x, y) | \quad \|x\|_2 \leq y\} \qquad \text{and} \qquad \hat{C} = \{(x, y) | \quad \|x\|_2^2 \leq y\}.$$

Determine whether the sets C and $\hat{C}$ are convex or not?

Solution: For this problem, I prove that both C and $\hat{C}$ are convex sets by showing that for points z_1 and $z_2 \in \mathbb{R}^{n+1}$ such that they are also in C and $\hat{C}$, then, for $\alpha \in (0, 1)$, $\alpha z_1 + (1-\alpha) z_2$ also belongs to C and $\hat{C}$, respectively.

Let $z_1 = \{\vec{x_1}, y_1\}$ and $z_2 = \{\vec{x_2}, y_2\}$, where $x_1, x_2 \in \mathbb{R}^n$. Then,

$$\alpha z_1 + (1-\alpha) z_2 = (\alpha \vec{x_1}, \alpha y_1) + ((1-\alpha)\vec{x_2}, (1-\alpha) y_2) = (\alpha \vec{x_1} + (1-\alpha)\vec{x_2}, \alpha y_1 + (1-\alpha) y_2).$$

Now, we need to show that this point satisfies the properties of sets C and $\hat{C}$. By triangle inequality, we have that

$$\|\alpha \vec{x_1} + (1-\alpha)\vec{x_2}\|_2 \leq \|\alpha \vec{x_1}\|_2 + \|(1-\alpha)\vec{x_2}\|_2. \tag{1}$$

However since $z_1, z_2 \in C$, we know that $\|\vec{x_1}\|_2 \leq y_1$ and $\|\vec{x_2}\|_2 \leq y_2$. Moreover, we know that $\alpha > 0$. Thus, (1) can be written as follows:

$$\|\alpha \vec{x_1}\|_2 + \|(1-\alpha)\vec{x_2}\|_2 = \alpha \|\vec{x_1}\|_2 + (1-\alpha)\|\vec{x_2}\|_2 \leq \alpha y_1 + (1-\alpha) y_2;$$

hence, C is a convex set.

Before proving the convexity of set $\hat{C}$, we derive an inequality that will be useful in our proof:

$$\begin{aligned}
&\underbrace{\alpha(1-\alpha)}_{>0} \underbrace{(\|\vec{x_1}\| - \|\vec{x_2}\|)^2}_{>0} > 0 \\
&2\alpha(1-\alpha)\|\vec{x_1}\|\|\vec{x_2}\| < \alpha(1-\alpha)\|\vec{x_1}\|^2 + \alpha(1-\alpha)\|\vec{x_2}\|^2. && (2)
\end{aligned}$$

Next, we have, by Cauchy Schwarz and triangle inequality, that

$$\begin{aligned}
\|\alpha \vec{x_1} + (1-\alpha)\vec{x_2}\|_2^2 &\leq \alpha^2 \|\vec{x_1}\|_2^2 + (1-\alpha)^2 \|\vec{x_2}\|_2^2 + 2(\alpha)(1-\alpha)\|\vec{x_1}\|_2 \|\vec{x_2}\|_2 \\
&\leq \alpha \|\vec{x_1}\|_2^2 + (1-\alpha)\|\vec{x_2}\|_2^2 + \alpha(1-\alpha)\|\vec{x_1}\|^2 + \alpha(1-\alpha)\|\vec{x_2}\|^2 \quad \text{by (2)} \\
&= \|\vec{x_1}\|^2(\alpha^2 + \alpha - \alpha^2) + \|\vec{x_2}\|^2(1 - 2\alpha + \alpha^2 + \alpha - \alpha^2) \\
&= \alpha \|\vec{x_1}\|^2 + (1-\alpha)\|\vec{x_2}\|^2 \\
&\leq \alpha y_1 + (1-\alpha) y_2.
\end{aligned}$$

Hence, we have proved that $\hat{C}$ is also a convex set. The last line holds, since if z_1, for example, is in $\hat{C}$, then $\|\vec{x}\|_2^2 \leq y$.

2. Consider the smooth (differentiable) functions $h : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$. Prove that the function

$$f = h \circ g : \mathbb{R}^n \to \mathbb{R}$$

where

$$f(x) = h(g(x)) \qquad \textit{and} \qquad \text{dom } f = \{x \in \text{dom } g | g(x) \in \text{dom } h\}$$

is convex if one of the following conditions on h and g holds.

(a) If h and g are convex functions, and h is nondecreasing, or

(b) if h is convex and nonincreasing, and g is concave.

To adhere to the formal definition of convex functions, we assume that the points x and y that we take, are not only in the $domf$, but also belong to a convex set. Assume that both h and g are convex functions. Then, by definition of convex functions, we have that for $x, y \in domf \in \mathbb{R}^n$ and $t \in (0,1)$,

$$g(tx + (1-t)y) \leq tg(x) + (1-t)g(y). \tag{3}$$

Since we know that h is non-decreasing, we have that for $a < b$, $h(a) \leq h(b)$. Applying (3) and using the fact that h is convex, we have that

$$f(tx + (1-t)y) = h(g(tx + (1-t)y)) \leq h(tg(x) + (1-t)g(y)) \leq th(g(x)) + (1-t)h(g(y)).$$

Next, let's assume that h is convex and non-increasing, and that g is concave. The latter indicates that $-g$ is convex. Then, for $x, y \in domf \in \mathbb{R}^n$,

$$g(tx + (1-t)y) = -(-g(tx + (1-t)y)) \geq tg(x) + (1-t)g(y).$$

Note that since $-g$ is convex, we have that $-g(tx+(1-t)y) \leq -tg(x) - (1-t)g(y)$, which, in turn, implies that $g(tx+(1-t)y) \geq tg(x)+(1-t)g(y)$. Next, we know that h is non-increasing, meaning that for $a < b$, $h(a) \geq h(b)$. Thus,

$$f(tx + (1-t)y) = \overbrace{h(\underbrace{g(tx + (1-t)y)}_{b>})}^{h(b)} \leq \overbrace{h(\underbrace{tg(x) + (1-t)g(y)}_{>a})}^{h(a)} \leq \underbrace{th(g(x)) + (1-t)h(g(y))}_{\text{since } h \text{ is convex}}.$$