Problem 4 *(subdifferential image)*

Prove or disprove by counterexample: Let $\psi : \mathbb{R}^n \to \mathbb{R}$ be a convex function and let $S \subset \mathbb{R}^n$ be an open convex set. Then $\partial\phi(S) \subset \mathbb{R}^n$ is a convex set.

Disprove by counterexample

The statement is false. For an easy counterexample let $\phi : \mathbb{R}^n \to \mathbb{R}$ be defined by $\phi(x) = |x|$. Convexity of ϕ can be seen by triangle inequality. For $x \neq 0$ we know that ϕ is differentiable with $\nabla\phi(x) = \frac{x}{|x|}$. If $x = 0$ by definition of the subdifferential

$$\partial\phi(0) = \{y \in \mathbb{R}^n | \phi(z) \geq \phi(0) + y \cdot (z - 0) \quad \forall z \in \mathbb{R}^n\} \tag{1}$$

If $y \in \partial\phi(0)$ then we can choose in particular $z = y$ and see $|y| \geq y \cdot y = |y|^2$. Dividing by $|y|$ this means $|y| \leq 1$. Hence $\overline{B_1(0)} \subseteq \partial\phi(0)$. On the other hand for $y \in \overline{B_1(0)}$ it follows by Cauchy-Schwartz inequality $y \cdot z \leq |y| + |z| = |z|$ and thus $y \in \partial\phi(0)$. Together $\partial\phi(0) = \overline{B_1(0)}$ and thus

$$\partial\phi(x) = \begin{cases} \overline{B_1(0)}, & x = 0, \\ \{\frac{x}{|x|}\}, & x \neq 0. \end{cases}$$

Now let $S = \{x \in \mathbb{R}^n | x_1 > 0\}$. This set is open and convex. We conclude

$$\partial\phi(S) = \left\{x \in \mathbb{R}^n \middle| |x| = 1, x_1 > 0\right\}, \tag{2}$$

which is clearly not a convex set (easiest case is $n = 2$ but we don't need this restriction).

$\square$

Problem 7 *(c-Legendre transform and c-subdifferential)*

Let $c(x,y) = \frac{|x-y|^2}{2}$ be defined on $M^+ \times M^- = \mathbb{R}^n \times \mathbb{R}^n$. Consider the c-convex function $\phi(x) = \max\{-\frac{|x-y_0|^2}{2}, -\frac{|x-y_1|^2}{2}\}$ on $\mathbb{R}^n$ for a fixed pair $y_0, y_1 \in \mathbb{R}^n$ with $y_0 \neq y_1$.

i) Compute $\partial^c \phi$.

ii) Compute the c^*-Legendre transform ϕ^{c^*}.

iii) Compute $\partial^{c^*} \phi^{c^*}$.

i) Compute $\partial^c \phi$.

Recap the definition of the c-subdifferential,

$$\partial^c \phi(x) = \{y \in M^- | \phi(z) \geq \phi(x) - c(z,y) + c(x,y) \quad \forall z \in M^+\}. \tag{3}$$

Since $|x-y|^2 = |x|^2 - 2x \cdot y + |y|^2$ we can conclude

$$\begin{aligned}
\partial^c \phi(x) &= \left\{ y \in \mathbb{R}^n \middle| \phi(z) \geq \phi(x) - \frac{|z-y|^2}{2} + \frac{|x-y|^2}{2} \quad \forall z \in \mathbb{R}^n \right\} && (4) \\
&= \left\{ y \in \mathbb{R}^n \middle| \phi(z) \geq \phi(x) - \frac{|z|^2}{2} + z \cdot y + \frac{|x|^2}{2} - x \cdot y \quad \forall z \in \mathbb{R}^n \right\} && (5) \\
&= \left\{ y \in \mathbb{R}^n \middle| \phi(z) + \frac{|z|^2}{2} \geq \phi(x) + \frac{|x|^2}{2} + y \cdot (z - x) \quad \forall z \in \mathbb{R}^n \right\} = \partial \hat{\phi}(x), && (6)
\end{aligned}$$

where $\hat{\phi}(x) = \phi(x) + \frac{|x|^2}{2}$. So we found a nice way to compute the c-subdifferential for this special cost function. Consider the three disjoint cases: A) $|x-y_0| < |x-y_1|$, B) $|x-y_0| > |x-y_1|$ and C) $|x-y_0| = |x-y_1|$.

A) It is $\hat{\phi}(x) = -\frac{|x-y_0|^2}{2} + \frac{|x|^2}{2} = x \cdot y_0 - \frac{|y_0|^2}{2}$, which is differentiable with $\nabla \hat{\phi}(x) = y_0$.

B) It is $\hat{\phi}(x) = -\frac{|x-y_1|^2}{2} + \frac{|x|^2}{2} = x \cdot y_1 - \frac{|y_1|^2}{2}$, which is differentiable with $\nabla \hat{\phi}(x) = y_1$.

C) In *Villani, Topics in Optimal Transportation, page 54* we can find a theorem which states

$$\partial \hat{\phi}(x) = \overline{\text{Conv}} \left(\lim_{x_k \to x} \nabla \hat{\phi}(x_k) \right). \tag{7}$$

Using this and the previous cases we see immediately that

$$\partial \hat{\phi}(x) = \overline{\text{Conv}} \left(\{y_0\}, \{y_1\}\right) = [y_0, y_1]. \tag{8}$$

Where $[y_0, y_1]$ denotes the line segment between y_0 and y_1.

This leads us to the conclusion

$$\partial^c \phi(x) = \partial \hat{\phi}(x) = \begin{cases} \{y_0\}, & |x-y_0| < |x-y_1|, \\ [y_0, y_1], & |x-y_0| = |x-y_1|, \\ \{y_1\}, & |x-y_0| > |x-y_1|. \end{cases}$$

and in the general form

$$\partial^c \phi = \left\{(x, y_0) \middle| |x-y_0| < |x-y_1|\right\} \cup \left\{(x, y_1) \middle| |x-y_0| > |x-y_1|\right\} \cup \left\{\{x\} \times [y_0, y_1] \middle| |x-y_0| = |x-y_1|\right\}. \tag{9}$$

ii) Compute the c^*−Legendre transform ϕ^{c^*}.

Using the definition of ϕ we see that

$$\begin{aligned}
\phi^{c^*}(y) &= \sup_{x\in\mathbb{R}^n}\left[-\frac{|x-y|^2}{2}-\max\{-\frac{|x-y_0|^2}{2},-\frac{|x-y_1|^2}{2}\}\right] && (10)\\
&= \sup_{x\in\mathbb{R}^n}\left[-\frac{|x-y|^2}{2}+\min\{\frac{|x-y_0|^2}{2},\frac{|x-y_1|^2}{2}\}\right]. && (11)
\end{aligned}$$

Now we see immediately, that for $y_i, i = 1, 2$

$$\begin{aligned}
\phi^{c^*}(y_i) &= \sup_{x\in\mathbb{R}^n}\left[-\frac{|x-y_i|^2}{2}+\min\{\frac{|x-y_0|^2}{2},\frac{|x-y_1|^2}{2}\}\right] && (12)\\
&\leq \sup_{x\in\mathbb{R}^n}\left[-\frac{|x-y_i|^2}{2}+\frac{|x-y_i|^2}{2}\right] && (13)\\
&= 0, && (14)
\end{aligned}$$

and on the other hand

$$\phi^{c^*}(y_i) \geq \sup_{x\in\mathbb{R}^n}\left[-\frac{|x-y_i|^2}{2}\right] = 0. \tag{15}$$

To find more values of ϕ^{c^*} we can use c-convexity of ϕ from the task and the result of problem 6 part (3), which states that for a c-convex function it holds

$$y \in \partial^c\phi(x) \Leftrightarrow \phi^{c^*}(y) = -c(x,y) - \phi(x). \tag{16}$$

Actually that would have saved us the inequalities before for the cases $y = y_i, i = 1, 2$. Anyways, looking at $\partial^c\phi$ that we computed in i) we see that for $y \in [y_0, y_1]$

$$y \in \partial^c\phi(x) \Leftrightarrow \phi^{c^*}(y) = -c(x,y) - \phi(x) \quad \forall x \text{ statisfying } |x-y_0| = |x-y_1|. \tag{17}$$

So let us take $x = \frac{y_0+y_1}{2}$. Then

$$\begin{aligned}
\phi^{c^*}(y) &= \frac{1}{2}\left(-\left|\frac{y_0+y_1}{2}-y\right|^2+\left|\frac{y_0+y_1}{2}-y_0\right|^2\right) && (18)\\
&= \frac{1}{8}\left(|y_1-y_0|^2-|y_0+y_1-2y|^2\right) && (19)\\
&= \frac{1}{8}\left(|y_1|^2+|y_0|^2-2y_1\cdot y_0-|y_0|^2-|y_1|^2+2y_0\cdot y_1-4|y|^2+4(y_0+y_1)\cdot y\right) && (20)\\
&= -\frac{1}{2}\left(|y|^2+y_0\cdot y_1+-y_0\cdot y-y_1\cdot y\right) && (21)\\
&= -\frac{1}{2}(y-y_0)\cdot(y-y_1). && (22)
\end{aligned}$$

We have to figure out what happens in the remaining case where $y \notin [y_0, y_1]$. Let us further modify the expression for ϕ^{c^*} that we found in the beginning.

$$\begin{aligned}
\phi^{c^*}(y) &= \sup_{x\in\mathbb{R}^n}\left[-\frac{|x-y|^2}{2}+\min\{\frac{|x-y_0|^2}{2},\frac{|x-y_1|^2}{2}\}\right] && (23)\\
&= \frac{1}{2}\sup_{x\in\mathbb{R}^n}\left[-|x|^2-|y|^2+2x\cdot y+\min\left\{|x|^2+|y_0|^2-2x\cdot y_0,|x|^2+|y_1|^2-2x\cdot y_1\right\}\right] && (24)\\
&= \frac{-|y|^2}{2}+\sup_{x\in\mathbb{R}^n}\left[\min\left\{\frac{-|y_0|^2}{2}+x\cdot(y-y_0),\frac{-|y_1|^2}{2}+x\cdot(y-y_1)\right\}\right] && (25)
\end{aligned}$$

If we let now, for fixed x, be α_0 be the angle between x and $y - y_0$ and α_1 the one between x and $y - y_1$, then we can write

$$\phi^{c^*}(y) = \frac{-|y|^2}{2} + \sup_{x\in\mathbb{R}^n} \left[\min\left\{\frac{-|y_0|^2}{2} + |x||y-y_0|\cos(\alpha_0), \frac{-|y_1|^2}{2} + |x||y-y_1|\cos(\alpha_1)\right\}\right]. \quad (26)$$

From this we can see that $\phi^{c^*}(y)$ is infinite! Indeed, we find an $x \in \mathbb{R}^n$ such that $0 \leq \alpha_0, \alpha_1 < \frac{\pi}{2}$ since the angle between $y - y_0$ and $y - y_1$ has to be strictly less than π. So both cosines will be fixed numbers bigger than zero. Scaling our x will not influence the angles but will increase the absolute value. Thus considering nx and letting $n \to \infty$ gives the desired statement.

iii) Compute $\partial^{c^*}\phi^{c^*}$

In the task it is already given, that ϕ is a c-convex function. We can use the result of problem 6 part (4) which states that for a c-convex function $y \in \partial^c\phi(x)$ iff $x \in \partial^{c^*}\phi^{c^*}(y)$. As seen before

$$\partial^c\phi(x) = \begin{cases} \{y_0\}, & |x-y_0| < |x-y_1|, \\ [y_0, y_1], & |x-y_0| = |x-y_1|, \\ \{y_1\}, & |x-y_0| > |x-y_1|. \end{cases}$$

Hence

$$\partial^{c^*}\phi^{c*}(y) = \begin{cases} \{x \in \mathbb{R}^n \big| |x-y_0| < |x-y_1|\}, & y = y_0, \\ \{x \in \mathbb{R}^n \big| |x-y_0| = |x-y_1|\}, & y \in [y_0, y_1], \\ \{x \in \mathbb{R}^n \big| |x-y_0| > |x-y_1|\}, & y = y_1. \end{cases}$$

And in general form

$$\partial^{c^*}\phi^{c*} = \big\{(y_0, x)\big||x-y_0| < |x-y_1|\big\} \cup \big\{(y_1, x)\big||x-y_0| > |x-y_1|\big\} \cup \big\{[y_0, y_1] \times \{x\}\big||x-y_0| = |x-y_1|\big\}. \quad (27)$$

$\square$

Problem 8 *(Legendre transform, subdifferential and c-subdifferential)*

Let $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ be given by $\varphi(x) = |x|$.

(4) Let $c(x,y) = \frac{|x-y|^2}{2}$ defined on $M^+ \times M^- = \mathbb{R}^n \times \mathbb{R}^n$. Compute $\partial^c \varphi$.

Analogously to problem 7, recap the definition of the c-subdifferential,

$$\partial^c \phi(x) = \{y \in M^- | \phi(z) \geq \phi(x) - c(z,y) + c(x,y) \quad \forall z \in M^+\}. \tag{28}$$

Since $|x-y|^2 = |x|^2 - 2x \cdot y + |y|^2$ we can conclude

$$\begin{aligned}
\partial^c \phi(x) &= \left\{ y \in \mathbb{R}^n \middle| \phi(z) \geq \phi(x) - \frac{|z-y|^2}{2} + \frac{|x-y|^2}{2} \quad \forall z \in \mathbb{R}^n \right\} && (29)\\
&= \left\{ y \in \mathbb{R}^n \middle| \phi(z) \geq \phi(x) - \frac{|z|^2}{2} + z \cdot y + \frac{|x|^2}{2} - x \cdot y \quad \forall z \in \mathbb{R}^n \right\} && (30)\\
&= \left\{ y \in \mathbb{R}^n \middle| \phi(z) + \frac{|z|^2}{2} \geq \phi(x) + \frac{|x|^2}{2} + y \cdot (z-x) \quad \forall z \in \mathbb{R}^n \right\} = \partial \hat{\phi}(x), && (31)
\end{aligned}$$

where $\hat{\phi}(x) = \phi(x) + \frac{|x|^2}{2}$. So it suffices to compute $\partial \hat{\phi}$ which is easier. For $x \neq 0$ the function $\hat{\phi}$ is differentiable with

$$\nabla \hat{\phi}(x) = \frac{x}{|x|} + x. \tag{32}$$

The more complicated case is when $x = 0$ because of the non-differentiability of $\hat{\phi}$. Luckyly the previously mentioned theorem of Villani will help us.

$$\partial \hat{\phi}(0) = \overline{\text{Conv}} \left(\lim_{x_k \to 0} \nabla \hat{\phi}(x_k) \right) = \overline{\text{Conv}} \left(\lim_{x_k \to 0} \nabla \left(|x_k| + \frac{|x_k|^2}{2} \right) \right) = \overline{\text{Conv}} \left(\lim_{x_k \to 0} \nabla (|x_k|) + \nabla \left(\frac{|x_k|^2}{2} \right) \right). \tag{33}$$

Now use that $x_k \neq 0$ and compute the gradients explicitly.

$$\partial \hat{\phi}(0) = \overline{\text{Conv}} \left(\lim_{x_k \to 0} \frac{x_k}{|x_k|} + x_k \right) = \overline{\text{Conv}} \left(S_1(0) \right) = \overline{B_1(0)}, \tag{34}$$

where $S_1(0)$ denotes the unit sphere and $B_1(0)$ the unit ball. All together

$$\partial^c \phi(x) = \partial \hat{\phi}(x) \begin{cases} \overline{B_1(0)}, & x = 0, \\ \frac{x}{|x|} + x, & x \neq 0. \end{cases}$$

And in general form

$$\partial^c \phi = \left\{ (x, \frac{x}{|x|} + x) \big| x \in \mathbb{R}^n - \{0\} \right\} \cup \left\{ \{0\} \times \overline{B_1(0)}. \right\} \tag{35}$$

(5) Let $c(x,y) = |x - y|$ defined on $M^+ \times M^- = \mathbb{R}^n \times \mathbb{R}^n$. Compute $\partial^c \varphi$.

By definition

$$\partial^c \varphi(x) = \{y \in \mathbb{R}^n | \varphi(z) \geq \varphi(x) - c(z,y) + c(x,y) \quad \forall z \in \mathbb{R}^n\}. \tag{36}$$

In particular for fixed $y \in \partial^c \varphi(x)$ with $z = 0$ we have

$$
\begin{aligned}
0 &\geq |x| - |-y| + |x-y| && (37)\\
|y| - |x| &\geq |x-y|. && (38)
\end{aligned}
$$

Note that as a consequence of the triangle inequality it holds

$$
|y| = |(y-x)+x| \leq |y-x| + |x| \Rightarrow |y| - |x| \leq |x-y|, \tag{39}
$$

which implies equality in our special case. Equality in the triangle inequality holds iff one of the following conditions is statisfied

i) $x = 0$.

ii) $x - y = 0$.

iii) $x = \lambda(y-x)$ for some $\lambda \geq 0$(vectors are parallel).

The last two cases can be combined to $y = \lambda x$ for some $\lambda \geq 1$. Let us show that the two remaining conditions are sufficient for $y \in \partial^c(x)$. Indeed, if $x = 0$ we stay with

$$
|z| \geq -|z-y| + |y| \quad \forall z \in \mathbb{R}^n \tag{40}
$$

which is true by the triangle inequality. If $y = \lambda x$ for some $\lambda \geq 1$, then again using triangle inequality we get

$$
\begin{aligned}
|z| &\geq |y| - |z-y| && (41)\\
&= \lambda|x| - |z-y| && (42)\\
&= |x| - |z-y| + |(\lambda-1)x| && (43)\\
&= |x| - |z-y| + |y-x|. && (44)
\end{aligned}
$$

and thus $\lambda x \in \partial^c \varphi(x)$ for $\lambda \geq 1$. Hence

$$
\partial^c \varphi(x) = \begin{cases} \{\lambda x \big| \lambda \geq 1\}, & x \neq 0, \\ \mathbb{R}^n, & x = 0. \end{cases}
$$

And in general form

$$
\partial^c \varphi = \Big\{(x, \lambda x)\Big| x \in \mathbb{R}^n - \{0\}, \lambda \geq 1\Big\} \cup \{\{0\} \times \mathbb{R}^n\} \tag{45}
$$

$\square$