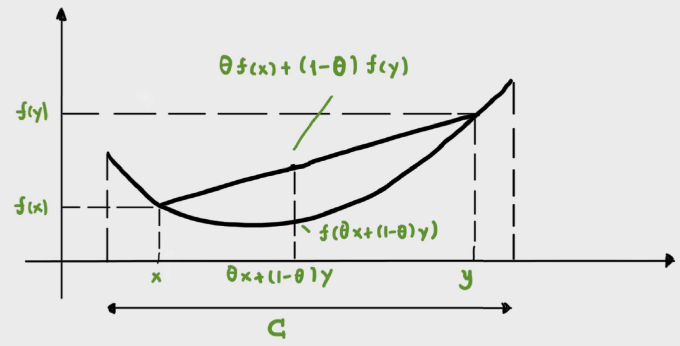
B. Convex Functions - Summary of results

We define a **real-valued** convex function

Definition Let $C \subseteq \mathbb{R}^n$ be a convex set. We say that a function $f: C \to \mathbb{R}$ is **convex** if

$$f(\theta x + (1-\theta) y) \leq \theta f(x) + (1-\theta) f(y) \quad \forall x, y \in C, \forall \theta \in [0,1]$$

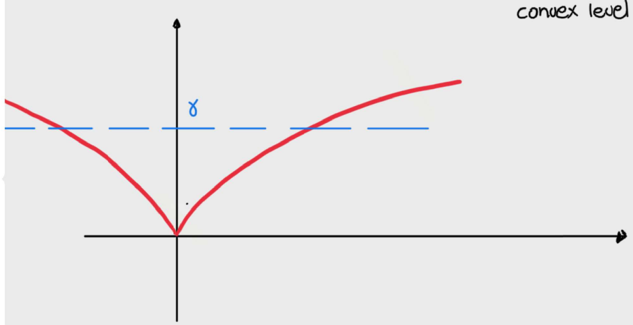


The linear interpolation $\theta f(x) + (1-\theta) f(y)$ **overestimates** the function value $f(\theta x + (1-\theta) y)$ for all $\theta \in [0,1]$.

A function $f: C \to \mathbb{R}$, where $C \subseteq \mathbb{R}^n$ is convex is called **concave** if the function $(-f)$ is convex.

Lemma: If $f: C \to \mathbb{R}$ is a function and γ is a scalar, the sets $\{x \in C: f(x) \leq \gamma\}$ and $\{x \in C: f(x) < \gamma\}$, are called **sublevel sets** of f. If f is a convex function, then all its level sets are convex

The converse is not true, e.g., $f(x) = \sqrt{|x|}$ nonconvex with convex level sets



Extended Real-Valued Functions

We extend the domain to the whole $\mathbb{R}^n$, but allow f to take infinite values and introduce the notion of **extended real-valued functions**.

Motivation

1. we will encounter operations resulting in functions that can take infinite values, e.g., the function $f(x) = \sup_{i \in I} f_i(x)$, where I infinite index or the conjugate of a real-valued function often take infinite values

2. convex functions over a convex set C, that cannot be extended to a convex function over the whole $\mathbb{R}^n$.

Definition: The **epigraph** of a function $f: \mathbb{R}^n \to (-\infty, +\infty]$ is defined by

$$\text{epi}(f) := \{(x, w) \in \mathbb{R}^n \times \mathbb{R} : f(x) \leq w\} \subseteq \mathbb{R}^{n+1}$$

The **effective domain** of f is

$$\mathrm{dom}(f) = \{x \in \mathbb{R}^n : f(x) < +\infty\} = \{x \in \mathbb{R}^n : \exists\, w \in \mathbb{R} \text{ s.t. } (x,w) \in \mathrm{epi} f\}$$

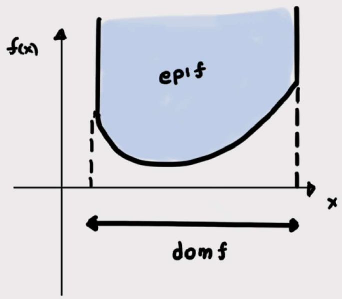
We will always assume that f is **proper**, i.e., $\mathrm{dom} f \neq \phi$.
(equivalently the epigraph of f is nonempty)

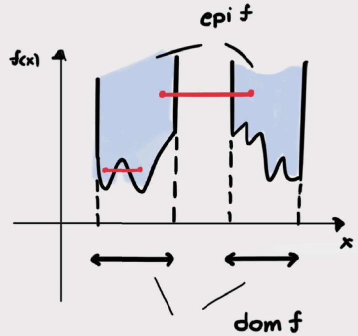
A **proper** function $f: \mathbb{R}^n \to (-\infty, +\infty]$ is **convex** if and only if

$$f(\theta x + (1-\theta) y) \leq \theta f(x) + (1-\theta) f(y), \quad \text{for all } x, y \in \mathbb{R}^n, \ \theta \in [0,1]$$

(definition is consistent with previous one. Moreover implies that $\mathrm{dom} f$ is convex)

We have f convex $\Leftrightarrow$ $\mathrm{epi} f$ is convex set in $\mathbb{R}^{n+1}$





$f: \mathbb{R}^n \to (-\infty, +\infty]$ is called **closed** if $\mathrm{epi} f$ is a closed set in $\mathbb{R}^{n+1}$!

Functional Operations Preserving Convexity

a. Positive Combinations

If $f_j : \mathbb{R}^n \to (-\infty, +\infty]$ proper and convex $\forall j = 1, \ldots, m$ and $t_j > 0$ for $j = 1, \ldots, m$. Then, under the assumption $\bigcap_{i \in I} \mathrm{dom} f_i \neq \phi$, $\sum_{i=1}^{m} t_i f_i$ is convex and proper.

b. Pointwise Supremum

$$f(x) = \sup_{i \in I} f_i(x) \quad \text{is convex}$$

(convex function $\forall i \in I$) (Assumption $\mathrm{dom} f \neq \phi$)

since $\mathrm{epi} f = \bigcap_{i \in I} \mathrm{epi} f_i$ (convex set $\forall i \in I$)

c. Partial Minimization

Minimizing over z a function that is jointly convex in two vectors x and z (Assumption: the infimum is always $> -\infty$)

d. Precomposition with an Affine map

$f: \mathbb{R}^n \to (-\infty, +\infty]$ convex $\Rightarrow$ $f(Ax+b)$ convex
(Assumption $\mathrm{dom} f \cap$ Image of affine map $\neq \phi$)

e. Postcomposition with an increasing convex function

The composite function $(h \circ g)(x) = h(g(x))$ is convex if one of the following holds ($g: \mathbb{R}^n \to \mathbb{R}$, $h: \mathbb{R}^n \to (-\infty, +\infty]$)

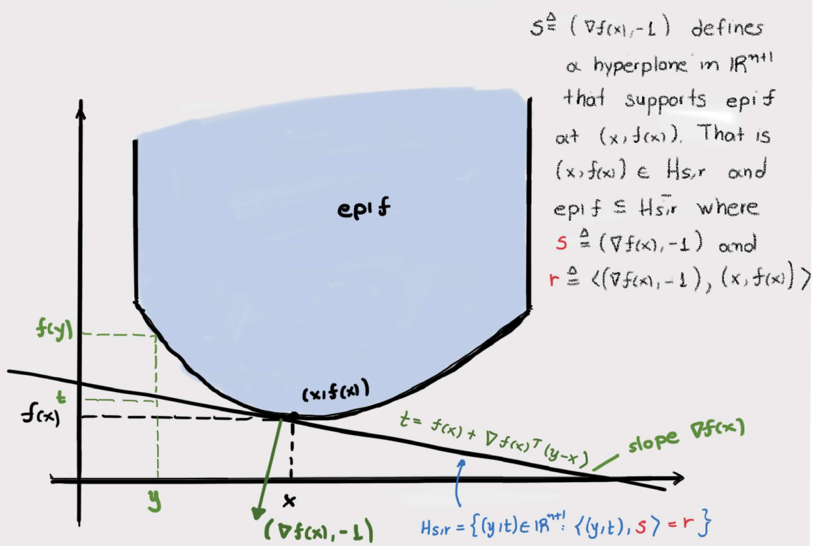
- h is convex increasing and g is convex
- h is convex decreasing and g is concave
- h is convex and g is affine

Similar reasoning for vector decomposition (Slide 18)

Differentiable Convex functions and Subdifferentias

Theorem (Convexity of differentiable functions) If $f: \mathbb{R}^n \to \mathbb{R}$ is differentiable, then f is convex if and only if

$$f(y) \geq f(x) + \nabla f(x)^T (y-x) \quad , \text{for all } x, y \in \mathbb{R}^n$$



The previous theorem tells us that if f is differentiable, then f is convex if and only if $(\nabla f(x), -1)$ is a (non-vertical) supporting hyperplane of epi f at the boundary point $(x, f(x))$. and epi f is contained in the corresponding (negative) halfspace. **So we get an affine global underestimator of f.**

Definition: If f is proper and convex, then $\forall x \in \text{dom} f$ we define the **subdifferential**

$$\partial f(x) = \{ g \in \mathbb{R}^n : f(y) \geq f(x) + \langle g, y-x \rangle \ \forall y \in \mathbb{R}^n \}$$

Clearly $g \in \partial f(x) \iff (g, -1)$ is a supporting hyperplane (non-vertical) of epi f at $(x, f(x))$ and epi f is contained in the corresponding negative halfspace.

Recall definition of normal cone from Lecture 1, then we can write

$$g \in \partial f(x) \iff (g, -1) \in N_{\text{epi} f}(x, f(x))$$

Facts: Assume that $x \in \text{int}(\text{dom} f)$ and that f is closed, proper and convex. Then,

(1) $\partial f(x) \neq \phi$ convex, closed and bounded

(2) If f is differentiable at x, then $\partial f(x) = \{ \nabla f(x) \}$

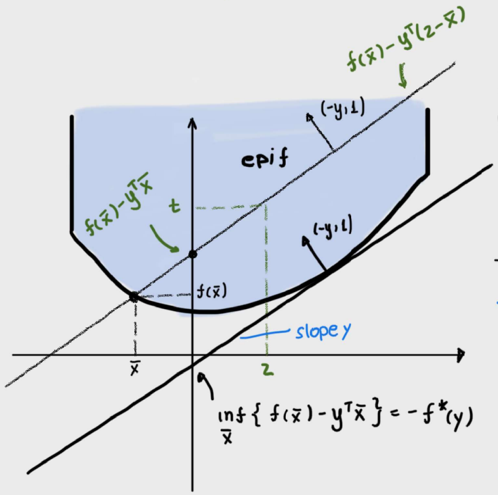
Theorem (Second order condition for convexity) A twice differentiable function $f: \mathbb{R}^n \to \mathbb{R}$ is convex if and only if $\nabla^2 f(x) \succeq 0$ for all $x \in \mathbb{R}^n$

Conjugate functions

Definition: The **conjugate function** of a function $f: \mathbb{R}^n \to (-\infty, +\infty]$

$$f^*(y) = \sup_{x \in \mathbb{R}^n} \{ x^T y - f(x) \}, \quad y \in \mathbb{R}^n$$

Geometrical interpretation of the definition:



Fix $y \in \mathbb{R}^n$. The crossing point of the vertical axis with the hyperplane that has normal $(-y,1)$ and passes through a point $(\bar{x}, f(\bar{x}))$ of the graph is $f(\bar{x}) - y^T\bar{x}$.

Thus the crossing point of the vertical axis with the hyperplane with normal $(-y,1)$ that **supports** the epigraph of f is

$$\inf_{\bar{x}} \{ f(\bar{x}) - y^T\bar{x} \} = -f^*(y).$$

f is closed and convex $\iff f^{**} = f$

biconjugate

Solved examples

Example 1 A proper function $f: \mathbb{R}^n \to (-\infty, +\infty]$ is **convex** if and only if: $\boldsymbol{f(\theta x + (1-\theta)y) \le \theta f(x) + (1-\theta) f(y)}$, $\forall\, x,y \in \mathbb{R}^n$, $\forall\, \theta \in [0,1]$ (*)
Show that f is convex if and only if epi f is a convex set in $\mathbb{R}^{n+1}$. In this case, show that dom f is a convex set in $\mathbb{R}^n$

Solution: We will show that epi f is convex $\iff$ (*) holds.

"$\Rightarrow$" Let $x, y \in \mathbb{R}^n$ and let $\theta \in [0,1]$. (*) holds trivially if $f(x) = +\infty$ or $f(y) = +\infty$. In any other case, $(x, f(x))$, $(y, f(y)) \in$ epi f which is convex. Thus

$$\theta\,(x, f(x)) + (1-\theta)\,(y, f(y)) \in \text{epi} f \iff$$
$$\iff (\theta x + (1-\theta) y,\ \theta f(x) + (1-\theta) f(y)) \in \text{epi} f \iff$$
$$\iff f(\theta x + (1-\theta) y) \le \theta f(x) + (1-\theta) f(y),$$

where the last equivalence holds by the definition of epi f

"$\Leftarrow$" Let $(x_1, t_1), (x_2, t_2) \in$ epi f. Then $f(x_1) \le t_1$ and $f(x_2) \le t_2$. In particular $x_1, x_2 \in$ dom f. Let $\theta \in [0,1]$, then

$$y \triangleq \theta\,(x_1, t_1) + (1-\theta)\,(x_2, t_2) = (\theta x_1 + (1-\theta) x_2,\ \theta t_1 + (1-\theta) t_2).$$

Therefore, $y \in$ epi $f \iff f(\theta x_1 + (1-\theta) x_2) \le \theta t_1 + (1-\theta) t_2$.
which holds, since (*) holds.

dom $f = \{ x \in \mathbb{R}^n : f(x) < +\infty \}$. If $x_1, x_2 \in$ dom f. Then $f(x_1), f(x_2) < +\infty$.
There for any $\theta \in [0,1]$, $f(\theta x_1 + (1-\theta) x_2) \underset{(*)}{\le} \theta f(x_1) + (1-\theta) f(x_2) < +\infty$
Thus $\theta x_1 + (1-\theta) x_2 \in$ dom f.

Example 2. The indicator function of a nonempty set C is convex if and only if $C \subseteq \mathbb{R}^n$ is convex.

Solution: $f(x) = I_C(x) = \begin{cases} 0 & \text{if } x \in C \\ \infty & \text{else} \end{cases}$ is convex if and only if epi f is convex. We have epi $f = C \times [0, +\infty)$ which is convex if and only if C is convex.

Example 3. Prove or disprove that the following functions $f : \mathbb{R}^n \to (-\infty, +\infty]$ are convex

1. $f_1(x,y) = xy$
2. $f_2(x_1, \dots, x_n) = \begin{cases} -(x_1 x_2 \dots x_n)^{1/n} & \text{if } x_1 > 0, \dots, x_n > 0 \\ +\infty & \text{otherwise} \end{cases}$

   $-f_2$ is called geometric mean and is concave since f_2 is convex
3. $f_3(x) = e^{\beta x^T A x}$, where $A \in S^n_{++}$ (positive definite symmetric $n \times n$ matrix) and $\beta > 0$.
4. $S_C(y) \overset{\Delta}{=} \sup_{x \in C} y^T x$, where $C \subseteq \mathbb{R}^n$ is a nonempty set (not necessarily convex)

   support function of C

   by definition $I_C^*(y) = \sup_{x \in \mathbb{R}^n} \{ y^T x - I_C(x) \} = \sup_{x \in C} \{ y^T x \} = S_C(y)$

Solution: 1. Not Convex. The function is twice continuously differentiable. We have $\nabla^2 f_1(x,y) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. This is not positive semidefinite (symmetric but eigenvalues $-1, 1$). Hence f_1 is nonconvex.

2. Convex. Denote $X = \text{dom} f_2 = \mathbb{R}^n_{++}$. It can be seen that f_2 is twice continuously differentiable over X (which is open) and its Hessian matrix is given by

$$\nabla^2 f_2(x) = \frac{f_2(x)}{n^2} \begin{bmatrix} \frac{1-n}{x_1^2} & \frac{1}{x_1 x_2} & \cdots & \frac{1}{x_1 x_n} \\ \vdots & & & \\ \frac{1}{x_n x_1} & \frac{1}{x_n x_2} & \cdots & \frac{1-n}{x_n^2} \end{bmatrix}$$

for all $x = (x_1, \dots, x_n) \in X$. From this, direct computation shows that for all $z = (z_1, \dots, z_n) \in \mathbb{R}^n$ and $x = (x_1, \dots, x_n) \in X$, we have

$$z^T \nabla^2 f_2(x) z = \frac{f_2(x)}{n^2} \left(\left(\sum_{i=1}^n \frac{z_i}{x_i} \right)^2 - n \sum_{i=1}^n \left(\frac{z_i}{x_i} \right)^2 \right) \geq 0,$$

since $f_2(x) < 0 \ \forall x \in X$, and for any real numbers $\alpha_1, \dots, \alpha_n$, we have

$$(\alpha_1 + \dots + \alpha_n)^2 \leq n(\alpha_1^2 + \dots + \alpha_n^2).$$ (prove it by induction in n)

Hence $\nabla^2 f_2(x) \succeq 0$, for all $x \in X$ (open), thus f_2 is convex

3. Convex. The function can be seen as a composition $h(g(x))$ of the function $h(t) = e^{\beta t}$ for $t \in \mathbb{R}$ and the function $g(x) = x^T A x$ for $x \in \mathbb{R}^n$. In this case, h is convex and monotonically increasing over $\mathbb{R}$, while g is convex over $\mathbb{R}^n$ (since $A \in S^n_{++}$). It follows that f_3 is convex

4. Convex. For a fixed x, the linear function $y \to y^T x$ is obviously convex. Therefore the support function is convex is convex as supremum of convex functions (regardless of whether C is convex or not). (exacty same reasoning holds for closedness).

Example 4. Determine if each set below is convex.

1. $\{ (x,y) \in \mathbb{R}^2_{++} \mid x/y \geq 1 \}$
2. $\{ (x,y) \in \mathbb{R}^2_+ \mid xy \leq 1 \}$

3. $\{ (x,y) \in \mathbb{R}^2_{++} \mid xy \geq 1 \}$

4. $\{ C \in S^n_{++} : \det C \geq (1/2)^n \}$

Solution: 1. **Convex**. The given set is $\{ (x,y) \in \mathbb{R}^2_{++} : x-y \geq 0 \}$
which is $\underbrace{\mathbb{R}^2_{++}}_{\text{convex}} \cap \underbrace{\{ (x,y) \in \mathbb{R}^2 : f(x,y) \geq 0 \}}_{\text{halfspace since } f \text{ is linear}}$, where $f(x,y) = x-y$

Intersection of convex sets is convex

2. **Nonconvex**. The points $(1/2, 2)$ and $(2, 1/2)$ are in the given set, but their average, $(5/4, 5/4)$ is not.
Important: As we saw in **Example 3**, the function $(x,y) \to xy$ is not convex. However, **we cannot conclude just from this** that its sublevel set $\{ (x,y) : xy \leq 1 \}$ is not convex.

3. **Convex**. The set can be written as $\{ (x,y) \in \mathbb{R}^2_{++} : \sqrt{xy} \geq 1 \}$ which is the 1-**superlevel set** of a **concave** function **(geometric mean)**
Thus, the set is convex

4. **Convex**. The set can be written as $\{ C \in S^n_{++} : -\log\det C \leq n\log 2 \}$. Thus the given set is the $(n\log 2)$ - **sublevel set** of the **convex** function $-\log\det C$ (on S^n_{++})
[proof, Boyd pg 74] **– Logdeterminant**

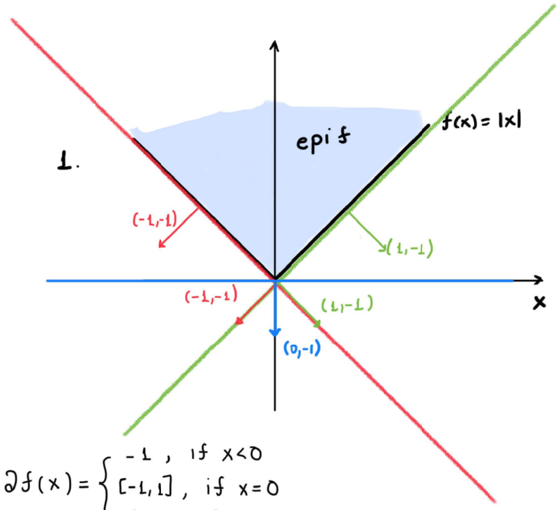
Example 5: Compute the **subdifferentials** of the following convex functions

1. $f(x) = |x|$

2. $f(x) = I_{[-1,1]}(x)$

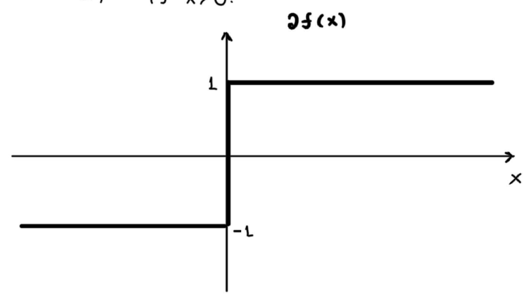
Solution:

1.

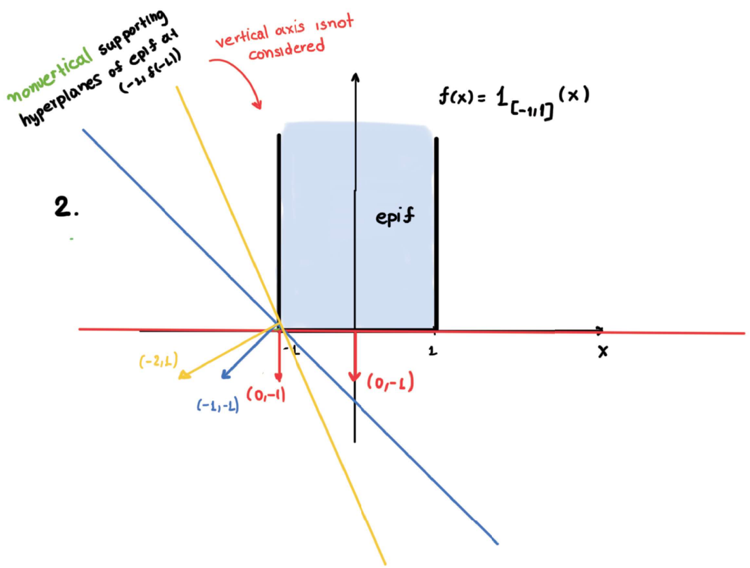


Reminder: $s \in \partial f(x)$ if and only if $(s, -1)$ defines a nonvertical supporting hyperplane of epi f at $(x, f(x))$

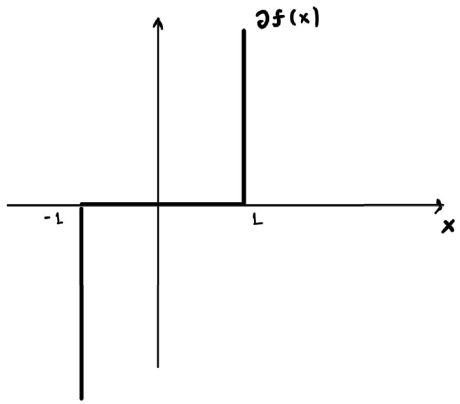
$$\partial f(x) = \begin{cases} -1, & \text{if } x<0 \\ [-1,1], & \text{if } x=0 \\ 1, & \text{if } x>0. \end{cases}$$



2.



$$\partial f(x) = \begin{cases} (-\infty, 0] & \text{if } x = -1 \\ 0 & \text{if } x \in (-1, 1) \\ [0, +\infty) & \text{if } x = 1 \\ \phi & \text{else} \end{cases}$$



Example 6. Compute the **conjugate functions** for the following convex functions

1. $f(x) = I_{[-1,1]}(x)$

2. $g(x) = |x|$

Solution. 1. We have $f^*(y) = \sup_{x \in [-1,1]} y\,x$

For $y \leq 0$, the maximizer is $x = -1$ and $f^*(y) = -y$
For $y \geq 0$, the maximizer is $x = 1$ and $f^*(y) = y$

Therefore $f^*(y) = |y|$.

2. Since $f(x) = I_{[-1,1]}(x)$ is (closed) convex $f^{**}(x) = f(x)$. Therefore, by previous question,
$\underset{\text{question 1}}{g^*(y) =} f^{**}(x) = f(x) = I_{[-1,1]}(x)$.

Example 7. Show that the **support function of a cone** $K \subseteq \mathbb{R}^n$ is the indicator function of its polar cone, i.e.,

$$\underbrace{S_K(y)}_{\text{Note: this is always equal to } I_K^*(y)} = I_{K^\circ}(y).$$

Solution: If $y \in K^\circ$, then $y^T x \leq 0$, for all $x \in K$. and for $x = 0$ we get $\langle y, x \rangle = 0$. Therefore,

$$S_K(y) = \max_{x \in K} y^T x = 0$$

If $y \notin K^{\circ}$, then there exists $\tilde{x} \in K$ such that $y^T \tilde{x} > 0$. Since, $\lambda \tilde{x} \in K$, for all $\lambda \geq 0$, it follows that

$$S_K(y) \geq y^T \lambda \tilde{x} = \lambda y^T \tilde{x} \quad \forall \ \lambda \geq 0.$$

Taking $\lambda \to +\infty$, we obtain that $S_K(y) = +\infty$ for $y \notin K^{\circ}$.