

# CONVEX ANALYSIS WORKSHOP

RYUTO SAITO

## 1. CONVEX FUNCTIONS

### 1.1. Definitions.

**Definition 1.1.1** (Convex function): Let  $S \subseteq \mathbb{R}^n$ . A function  $f : S \rightarrow \mathbb{R} \cup \{\infty\}$  is convex if  $\text{dom } f$  is a convex set and

$$\forall \mathbf{x}, \mathbf{y} \in \text{dom } f, \forall t \in [0, 1], f(t\mathbf{x} + (1-t)\mathbf{y}) \leq tf(\mathbf{x}) + (1-t)f(\mathbf{y}) \quad (1)$$

where  $\text{dom } f$  is the effective domain of  $f$ :

$$\text{dom } f := \{\mathbf{x} \in S \mid f(\mathbf{x}) < \infty\}. \quad (2)$$

**Definition 1.1.2** (Gradient): Let  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  be a differentiable function. The gradient of  $f$  at  $\mathbf{x} \in \mathbb{R}^n$ , denoted  $\nabla f(\mathbf{x})$ , is an  $n$ -dimensional vector whose entries are given by

$$(\nabla f(\mathbf{x}))_i := \frac{\partial f(\mathbf{x})}{\partial x_i}. \quad (3)$$

The gradient of  $f$  is the vector containing all the partial derivatives. Element  $i$  of the gradient is the partial derivative of  $f$  with respect to  $x_i$ .

### 1.2. Lemma.

**Lemma 1.2.1:** A differentiable function  $f : \mathbb{R} \rightarrow \mathbb{R}$  is convex if and only if

$$f(y) \geq f(x) + f'(x)(y - x) \quad (4)$$

for all  $x$  and  $y$  in  $\text{dom } f$ .

**Lemma 1.2.2:** Let  $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\infty\}, g : \mathbb{R} \rightarrow \mathbb{R} \cup \{\infty\}, \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, t \in [0, 1]$  and

$$g(t) = f(t\mathbf{y} + (1-t)\mathbf{x}). \quad (5)$$

Then,  $g$  is convex if and only if  $f$  is convex.

*Proof:* ( $\implies$ ) Let  $\theta \in [0, 1]$  and  $\mathbf{x}, \mathbf{y} \in \text{dom } f$ . For any  $t_1, t_2 \in \mathbb{R}$ ,

$$\begin{aligned}
& g(\theta t_1 + (1 - \theta)t_2) \\
&= f((\theta t_1 + (1 - \theta)t_2)\mathbf{y} + (1 - (\theta t_1 + (1 - \theta)t_2))\mathbf{x}) \\
&= f(\theta t_1 \mathbf{y} + (1 - \theta)t_2 \mathbf{y} + \mathbf{x} - \theta t_1 \mathbf{x} - (1 - \theta)t_2 \mathbf{x}) \\
&= f(\theta t_1 \mathbf{y} + \theta \mathbf{x} - \theta t_1 \mathbf{x} + (1 - \theta)t_2 \mathbf{y} + (1 - \theta)\mathbf{x} - (1 - \theta)t_2 \mathbf{x}) \quad (6) \\
&= f(\theta(t_1 \mathbf{y} + (1 - t_1)\mathbf{x}) + (1 - \theta)(t_2 \mathbf{y} + (1 - t_2)\mathbf{x})) \\
&\leq \theta f(t_1 \mathbf{y} + (1 - t_1)\mathbf{x}) + (1 - \theta)f(t_2 \mathbf{y} + (1 - t_2)\mathbf{x}) \\
&= \theta g(t_1) + (1 - \theta)g(t_2)
\end{aligned}$$

Thus,  $g$  is convex.

( $\Leftarrow$ ) Let  $\mathbf{x}, \mathbf{y} \in \mathbf{dom} f$  and  $t \in \mathbb{R}$ . For any  $\theta \in [0, 1]$ ,

$$\begin{aligned}
f(\theta \mathbf{y} + (1 - \theta)\mathbf{x}) &= g(\theta) \\
&= g(\theta \cdot 1 + (1 - \theta) \cdot 0) \\
&\leq \theta g(1) + (1 - \theta)g(0) \\
&= \theta f(\mathbf{y}) + (1 - \theta)f(\mathbf{x})
\end{aligned} \quad (7)$$

Thus,  $f$  is convex. □

### 1.3. Exercise.

**Proposition 1.3.1** (First-order convexity condition): Suppose  $f$  is differentiable. Then  $f$  is convex if and only if  $\mathbf{dom} f$  is convex and

$$\forall \mathbf{x}, \mathbf{y} \in \mathbf{dom} f, f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^\top (\mathbf{y} - \mathbf{x}). \quad (8)$$

*Proof:* Let  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ ,  $\mathbf{z} = t\mathbf{y} + (1 - t)\mathbf{x}$  for  $t \in [0, 1]$ , and

$$g(t) = f(t\mathbf{y} + (1 - t)\mathbf{x}). \quad (9)$$

Then, using chain rule,

$$\begin{aligned}
g'(t) &= \frac{d}{dt} f(t\mathbf{y} + (1-t)\mathbf{x}) \\
&= \frac{d}{dt} f(\mathbf{z}) \\
&= \sum_{i=1}^n \frac{d}{dt} z_i \frac{\partial}{\partial z_i} f(\mathbf{z}) \\
&= \left( \frac{\partial}{\partial z_1} f(\mathbf{z}), \dots, \frac{\partial}{\partial z_n} f(\mathbf{z}) \right) \begin{bmatrix} \frac{d}{dt} z_1 \\ \vdots \\ \frac{d}{dt} z_n \end{bmatrix} \\
&= \begin{bmatrix} \frac{\partial}{\partial z_1} f(\mathbf{z}) \\ \vdots \\ \frac{\partial}{\partial z_n} f(\mathbf{z}) \end{bmatrix}^T \begin{bmatrix} \frac{d}{dt} (ty_1 + (1-t)x_1) \\ \vdots \\ \frac{d}{dt} (ty_n + (1-t)x_n) \end{bmatrix} \\
&= \nabla f(\mathbf{z})^T \begin{bmatrix} y_1 - x_1 \\ \vdots \\ y_n - x_n \end{bmatrix} \\
&= \nabla f(t\mathbf{y} + (1-t)\mathbf{x})^T (\mathbf{y} - \mathbf{x}).
\end{aligned} \tag{10}$$

( $\implies$ ) Assume  $f$  is convex. From Lemma 1.2.2,  $g$  is convex. From Lemma 1.2.1, we have

$$g(1) \geq g(0) + g'(0), \tag{11}$$

which means that

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T (\mathbf{y} - \mathbf{x}). \tag{12}$$

( $\Leftarrow$ ) Assume that (8) holds for any  $\mathbf{x}, \mathbf{y} \in \text{dom } f$ . Let  $t_1, t_2 \in [0, 1]$ . Since  $\mathbf{x}, \mathbf{y} \in \text{dom } f$ ,  $t_1\mathbf{y} + (1-t_1)\mathbf{x} \in \text{dom } f$  and  $t_2\mathbf{y} + (1-t_2)\mathbf{x} \in \text{dom } f$ . From (8),

$$\begin{aligned}
&f(t_1\mathbf{y} + (1-t_1)\mathbf{x}) \\
&\geq f(t_2\mathbf{y} + (1-t_2)\mathbf{x}) \\
&\quad + \nabla f(t_2\mathbf{y} + (1-t_2)\mathbf{x})(t_1\mathbf{y} + (1-t_1)\mathbf{x} - (t_2\mathbf{y} + (1-t_2)\mathbf{x})) \\
&= f(t_2\mathbf{y} + (1-t_2)\mathbf{x}) \\
&\quad + \nabla f(t_2\mathbf{y} + (1-t_2)\mathbf{x})((t_1 - t_2)\mathbf{y} + (1-t_1 - (1-t_2))\mathbf{x}) \\
&= f(t_2\mathbf{y} + (1-t_2)\mathbf{x}) \\
&\quad + \nabla f(t_2\mathbf{y} + (1-t_2)\mathbf{x})((t_1 - t_2)\mathbf{y} - (t_1 - t_2)\mathbf{x}) \\
&= f(t_2\mathbf{y} + (1-t_2)\mathbf{x}) \\
&\quad + \nabla f(t_2\mathbf{y} + (1-t_2)\mathbf{x})(t_1 - t_2)(\mathbf{y} - \mathbf{x}).
\end{aligned} \tag{13}$$

That is

$$g(t_1) \geq g(t_2) + g'(t_2)(t_1 - t_2). \tag{14}$$

From Lemma 1.2.2, since  $g$  is convex, then  $f$  is convex.  $\square$

[1]

#### REFERENCES

1. Boyd, S., Vandenberghe, L.: Convex Optimization. Cambridge University Press (2004)

*Email address:* j2200071@gunma-u.ac.jp