CONVEX ANALYSIS WORKSHOP

RYUTO SAITO

1. Convex functions

I made this material referring to [1].

1.1. Definitions.

Definition 1.1.1 (Convex function): A function $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is convex if $\mathbf{dom}\, f$ is a convex set and

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

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

$$\mathbf{dom}\, f := \{\boldsymbol{x} \mid f(\boldsymbol{x}) < \infty\}. \tag{2}$$

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

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

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 $f : \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ is convex if and only if

$$f(y) \geq f(x) + f'(x)(y - x) \tag{4}$$

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

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

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

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

Proof: $(\Longrightarrow)$ Let $\theta \in [0,1]$. For any $t_1, t_2 \in \mathbf{dom}\, g$,

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

Thus, g is convex.

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

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

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

$$
\forall \boldsymbol{x}, \boldsymbol{y} \in \mathbf{dom}\, f, f(\boldsymbol{y}) \geq f(\boldsymbol{x}) + \nabla f(\boldsymbol{x})^{\mathrm{T}} (\boldsymbol{y} - \boldsymbol{x}). \tag{8}
$$

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

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

Then, using chain rule,

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

($\Longrightarrow$) 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

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

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

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

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. $\square$

References

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

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