Discussion 6

1. Convexity of Functions

Definition. A function $f\colon \mathbb{R}^n \to \mathbb{R}$ is convex if $\mathrm{dom}(f)$ is a convex set and if for all $\vec{x}, \vec{y} \in \mathrm{dom}(f)$ and $\theta \in [0, 1]$, we have,

$$f(\theta\vec{x} + (1-\theta)\vec{y}) \le \theta f(\vec{x}) + (1-\theta) f(\vec{y}). \tag{1}$$

The function f is strictly convex if the inequality is strict.

Definition. A function $f : \mathbb{R}^n \to \mathbb{R}$ is concave if $\mathrm{dom}(f)$ is a convex set and if for all $\vec{x}, \vec{y} \in \mathrm{dom}(f)$ and θ with $0 \le \theta \le 1$, we have,

$$f(\theta\vec{x} + (1-\theta)\vec{y}) \ge \theta f(\vec{x}) + (1-\theta) f(\vec{y}). \tag{2}$$

The function f is strictly concave if the inequality is strict.

Property. A function f is concave if and only if $-f$ is convex. An affine function is both convex and concave.

Property: Jensen's inequality. The inequality in Equation (1) is known as **Jensen's Inequality**. This can be extended to convex combinations of more than one point. If f is convex, and $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k \in \mathrm{dom}(f)$, and $\theta_1, \theta_2, \ldots, \theta_k \ge 0$ with $\sum_{i=1}^{k} \theta_i = 1$ then,

$$f(\theta_1\vec{x}_1 + \theta_2\vec{x}_2 + \cdots + \theta_k\vec{x}_k) \le \theta_1 f(\vec{x}_1) + \theta_2 f(\vec{x}_2) + \cdots + \theta_k f(\vec{x}_k). \tag{3}$$

Property: first order condition. Suppose f is differentiable. Then f is convex if and only if $\mathrm{dom}(f)$ is convex and

$$f(\vec{y}) \ge f(\vec{x}) + \nabla f(\vec{x})^\top (\vec{y} - \vec{x}), \tag{4}$$

for all $\vec{x}, \vec{y} \in \mathrm{dom}(f)$.

Property: Second order condition. Suppose f is twice differentiable. Then f is convex if and only if, $\mathrm{dom}(f)$ is convex and the Hessian of f, $\nabla^2 f(\vec{x})$, is positive semi-definite for all $\vec{x} \in \mathrm{dom}(f)$.

(a) **Restriction to a line**.

Show that a function f is convex if and only if for all $\vec{x} \in \mathrm{dom}(f)$ and all $\vec{v}$, the function $g : \mathrm{dom}(g) \to \mathbb{R}$ given by $g(t) = f(\vec{x} + t\vec{v})$ is convex for $\mathrm{dom}(g) = \{t \in \mathbb{R} \mid \vec{x} + t\vec{v} \in \mathrm{dom}(f)\}$.

(b) **Non-negative weighted sum**.

Show that the non-negative weighted sum of convex functions is convex: i.e. if $f_1, \ldots, f_n$ are n convex functions from $\mathbb{R}^n$ to $\mathbb{R}$ and $w_1, \ldots, w_n \in \mathbb{R}_+$ are n positive scalars, then the function:

$$f = \sum_{i=1}^{n} w_i f_i \tag{5}$$

is convex. To make the question easier, you can assume that the functions $f_1, \ldots, f_n$ are twice-differentiable.

(c) **Point-wise maximum**.

Show that if f_1 and f_2 are convex functions then their pointwise maximum f, defined by

$$f(\vec{x}) = \max(f_1(\vec{x}), f_2(\vec{x})), \tag{6}$$

with $\mathrm{dom}(f) = \mathrm{dom}(f_1) \cap \mathrm{dom}(f_2)$, is also convex.

2. Convexity of Constraint Sets

Let $f_1, \ldots, f_m, h_1, \ldots, h_p \colon \mathbb{R}^n \to \mathbb{R}$ be functions. Let $S \subseteq \mathbb{R}^n$ be defined as

$$S \doteq \left\{ \vec{x} \in \mathbb{R}^n \;\middle|\; \begin{array}{ll} f_i(\vec{x}) \leq 0 & \forall i = 1, \ldots, m \\ h_j(\vec{x}) = 0 & \forall j = 1, \ldots, p \end{array} \right\}. \tag{7}$$

Show that if $f_1, \ldots, f_m$ are convex functions, and $h_1, \ldots, h_p$ are affine functions, then S is a convex set.

3. Ridge Regression

Prove that the optimal solution to the ridge regression problem:

$$\min_{\vec{w} \in \mathbb{R}^p} \|X\vec{w} - \vec{y}\|_2^2 + \lambda \|\vec{w}\|_2^2, \tag{8}$$

where $X \in \mathbb{R}^{n \times p}$, $\lambda > 0$ and $\vec{y} \in \mathbb{R}^n$, is given by:

$$\vec{w}^* = (X^\top X + \lambda I)^{-1} X^\top \vec{y}. \tag{9}$$