Elementary Convexity

Let X be a linear space. Then,

Definition: A set $S \subset X$ is convex if whenever $x, y \in S$ then $\lambda x + (1-\lambda)y \in S, \forall \lambda \in (0,1)$.

Definition: Let $S \subset X$ be a convex set. Then, $f : S \to \mathbb{R}$ is said to be a convex function on S if for every $x, y \in S$

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y) \quad , \quad \forall \lambda \in (0,1).$$

f is said to be strictly convex on S if the inequality is strict $\forall x \neq y$.

In both the definitions of convex sets and functions, "convex" may be replaced by "concave" if the inequalities are reversed. Also, it is easily verified that sums of (strictly) convex functions are (strictly) convex.

If $f : S \to \mathbb{R}$, $S \subset \mathbb{R}^n$ is convex and f is sufficiently differentiable there are two important calculus-based equivalent definitions for convexity. First, suppose $x = (x_1, x_2, \ldots, x_n), y = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$. Whenever it exists, we define the n-dimensional gradient, $n \geq 1$, as

$$\nabla f(x) = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n} \right).$$

Similary, whenever it exists, we define the Hessian $H(x)$ of f as the matrix $H \in \mathbb{R}^{n \times n}$ where

$$H(x) = [H_{ij}] \quad , \quad H_{ij}(x) = \frac{\partial^2 f}{\partial x_i \partial x_j}$$

With these definitions we have the following two important results for $f : S \to \mathbb{R}$, $S \subset \mathbb{R}^n$:

Theorem: Let $f \in C^1$. Then f is (striclty) convex over S if and only if

$$f(y) - f(x) \geq \nabla f(x)(y - x) \quad , \quad \forall x, y \in S.$$

Theorem: Let $f \in C^2$. Then f is (strictly) convex over S if and only if the Hessian $H(x)$ is (strictly) positive definite $\forall x \in S$:

$$z^T H(x) z \geq 0 \quad , \quad \forall x \in S, \quad \forall z \neq 0$$

Also, recall Taylor's Theorem. If $f \in C^2$, then for each z there exists an $\alpha \in [0,1]$ such that

$$f(x + z) = f(x) + \nabla f(x) z + \frac{1}{2} z^T H(x + \alpha z) z$$