Proposition: Let $f: \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be an extended real-valued function. Then,

$$f \text{ is convex iff } \operatorname{epi}(f) \text{ is convex.}$$

(as a function) (as a set)

Moreover, let $S \subseteq \mathbb{R}^n$ be a set. Then,

$$S \text{ is convex iff } \mathbb{1}_S \text{ is convex.}$$

(Exercise)

Corollary: (Jensen's inequality)

Let f be as before. Then, f is convex iff

$$f\left(\sum_{k=1}^{l} \alpha_k x^k\right) \le \sum_{k=1}^{l} \alpha_k f(x^k)$$

for any $x^1, \dots, x^l \in \mathbb{R}^n$; $\alpha_1, \dots, \alpha_l \ge 0$ satisfying $\sum_{k=1}^{l} \alpha_k = 1$; $l \ge 2$.

Proof: ($\Leftarrow$) Obvious.

($\Rightarrow$) Suppose that f is convex. Note that

$$(x^k, f(x^k)) \in \operatorname{epi}(f) \quad \forall k. \qquad \operatorname{epi}(f) = \{(x,t) : f(x) \le t\}$$

By the proposition, $\operatorname{epi}(f)$ is convex.

$$\Rightarrow \quad \sum_{k=1}^{l} \alpha_k (x^k, f(x^k)) \in \operatorname{epi}(f)$$

$$\Rightarrow \quad \Big(\underbrace{\sum_{k=1}^{l} \alpha_k x^k}_{x}, \underbrace{\sum_{k=1}^{l} \alpha_k f(x^k)}_{t}\Big) \in \operatorname{epi}(f)$$

$$\Rightarrow \quad f\left(\sum_{k=1}^{l} \alpha_k x^k\right) \le \sum_{k=1}^{l} \alpha_k f(x^k).$$

Operations Preserving convexity

① (Non-negative Combinations) Let $f_1, \dots, f_m$ be convex functions, $\alpha_1, \dots, \alpha_m \ge 0$ be sclars. Then, $f = \sum_{k=1}^{m} \alpha_k f_k$ is convex.

$k=1$

Proof: Take any $x_1, x_2 \in \mathbb{R}^n$ and $\alpha \in [0,1]$

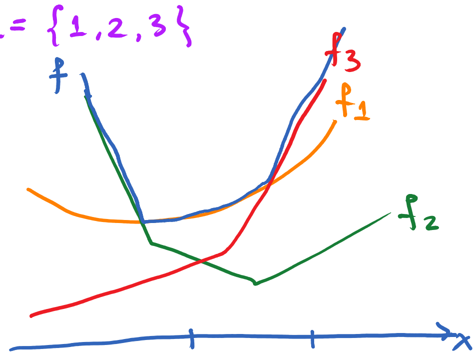
$$f(\alpha x_1 + (1-\alpha)x_2) = \sum_{k=1}^{m} \underbrace{\alpha_k}_{\geq 0} \underbrace{f_k(\alpha x_1 + (1-\alpha)x_2)}_{\leq \alpha f_k(x_1) + (1-\alpha) f_k(x_2) \text{ by assumption}}$$

$$\leq \sum_{k=1}^{m} \alpha_k \left[\alpha f_k(x_1) + (1-\alpha) f_k(x_2) \right]$$

$$= \alpha f(x_1) + (1-\alpha) f(x_2)$$

② (Pointwise Supremum) Let I be an index set and $\{f_i\}_{i\in I}$ is a family of convex functions. Then, $f = \sup_{i \in I} f_i$ is convex.

e.g.: $I = \{1,2,3\}$



e.g. Let $\sigma : \mathbb{R}^{m\times n} \to \mathbb{R}_+$ be defined as

$$\sigma(X) = \text{largest singular value of } X$$

$$= \sqrt{\text{largest eigenvalue of } \underbrace{XX^T}_{\text{psd}}}$$

Claim: $\sigma(\cdot)$ is convex.

Proof: By the Courant-Fischer theorem,

$$\sigma(X) = \max_{\substack{u \in \mathbb{R}^m, \|u\|_2 = 1 \\ v \in \mathbb{R}^n, \|v\|_2 = 1}} \underbrace{u^T X v}$$

Let $f_{u,v}(X) = u^T X v$, $I = \{(u,v) \in \mathbb{R}^m \times \mathbb{R}^n : \|u\|_2 = \|v\|_2 = 1\}$.

Then, $$\sigma(X) = \max_{(u,v)\in I} f_{u,v}(X)$$

Observe: $X \mapsto f_{u,v}(X)$ is linear! That is,

$f_{u,v}(\alpha X+\beta Y) = \alpha f_{u,v}(X) + \beta f_{u,v}(Y) \quad \forall X, Y;\ \alpha, \beta$. (check!)

Hence, $\sigma(\cdot)$ is convex by pointwise supremum principle.

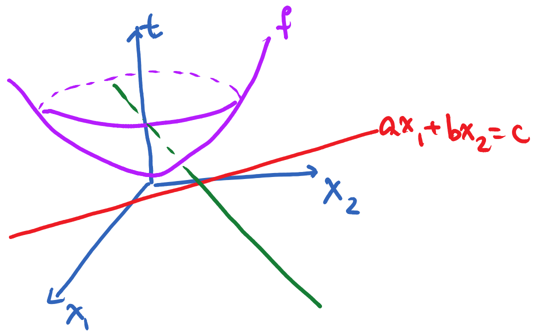
③ (Composition with Increasing Convex Functions)

Let $g: \mathbb{R}^n \to \mathbb{R}$ be convex, $h: \mathbb{R} \to \mathbb{R}$ be convex.

Suppose that h is increasing (i.e., $x \le y \Rightarrow h(x) \le h(y)$).

Then, $f = h \circ g$ (i.e., $f(x) = h(g(x))$) is convex.

④ (Restriction on Lines)



Definition: Given a point $x_0 \in \mathbb{R}^n$ and a direction $h \in \mathbb{R}^n \setminus \{0\}$, we call the set $\{x_0 + th : t \in \mathbb{R}\}$ a line through x_0 in the direction h.

Let f be a function. Define

$$\mathbb{R} \ni t \mapsto \tilde{f}_{x_0,h}(t) = f(\underbrace{x_0+th}_{\in \mathbb{R}^n})$$

to be the restriction of f along the line $\{x_0+th : t \in \mathbb{R}\}$.

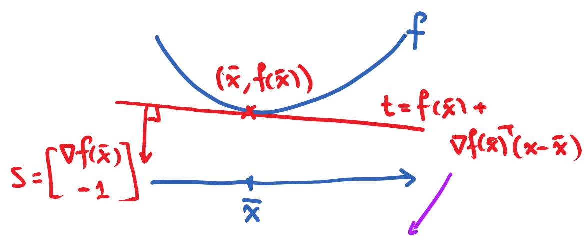
Then,

f is convex iff $\tilde{f}_{x_0,h}$ is convex for any $x_0 \in \mathbb{R}^n$, $h \in \mathbb{R}^n \setminus \{0\}$

Differentiable Convex Functions

Theorem: Let f be differentiable. Then, f is convex iff

$$f(x) \ge \underbrace{f(\bar{x}) + \nabla f(\bar{x})^T (x - \bar{x})}_{\text{affine function of } x} \quad \forall x, \bar{x} \in \mathbb{R}^n.$$



$$\begin{bmatrix} \nabla f(\bar{x}) \\ -1 \end{bmatrix}^T \begin{bmatrix} x \\ t \end{bmatrix} + f(\bar{x}) - \nabla f(\bar{x})^T \bar{x} = 0$$

Theorem: Let f be a twice continuously differentiable function (i.e., $\nabla f(x)$, $\nabla^2 f(x)$ exist and $\nabla^2 f(x)$ is continuous)

$$\left[\frac{\partial^2 f}{\partial x_i \partial x_j}\right]_{i,j} \qquad \overline{\mathbb{R}^n \to \mathbb{R}^{n\times n}}$$

Then, f is convex iff $\nabla^2 f(x) \succcurlyeq 0 \quad \forall x$