Theorem 2.6.1 [Elements of Information Theory; IInd Edition, Thomas M. Cover and Joy A. Thomas]

If the univariate function f has a second derivative that is non negative (positive) over an interval, the function is convex (strictly convex) over the interval.

Proof:

$$f''(x) \geq 0 \qquad x \in \underset{[a,b]}{I};$$

T.P.T f → is convex over I.

Let x_0 be a point in I.

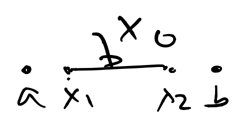
$$f(x) = f(x_0) + f'(x_0)(x-x_0) + \frac{1}{2} f''(\bar{x})(x-x_0)^2$$

where $\bar{x} \in [x_0, x]$

T.P.T $\boxed{f(\lambda x_1 + \overline{1-\lambda}\, x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2)}$ $\quad x_1 \in I$, $x_2 \in I$, $\lambda \in [0,1]$

Fix:

$X_0 = \lambda X_1 + \overline{1-\lambda}\, x_2$



Choose once, $x = x_1$ and next $x = x_2$.

$$\boxed{f(x) \geq f(x_0) + f'(x_0)(x - x_0)}$$

(1) $x = x_1$ $\quad x - x_0 = x_1(1-\lambda) - (1-\lambda)x_2$

$$x - x_0 = (1-\lambda)(x_1 - x_2)$$

$\lambda \times$ $\boxed{f(x_1) \geq f(\lambda x_1 + \overline{1-\lambda}\, x_2) + f'(x_0)(1-\lambda)(x_1 - x_2)}$

(2) Choose $x = x_2$.

$$x - x_0 = x_2 - (\lambda x_1 + \overline{1-\lambda}\, x_2)$$

$$= \lambda(x_2 - x_1) = -\lambda(x_1 - x_2)$$

$(1-\lambda)\times$ $\boxed{f(x_2) \geq f(\lambda x_1 + \overline{1-\lambda}\, x_2) - f'(x_0)\,\lambda\,(x_1 - x_2)}$

$$\lambda f(x_1) + (1-\lambda) f(x_2) \geq (\lambda + \overline{1-\lambda})\, f(\lambda x_1 + \overline{1-\lambda}\, x_2)$$

$$f(\lambda x_1 + \overline{1-\lambda}\, x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2)$$

$\therefore$ $f \Rightarrow$ is convex over the interval.

Ex 1: $f(x) = x^2$

$f''(x) = 2$ over $\mathbb{R}$.

$f(x)$ is strictly convex. $(-\infty, +\infty)$

Ex 2: $f(x) = e^x$

$f''(x) = e^x > 0 \quad \forall x.$

$\therefore$ e^x is also strictly convex.

Ex 3: $f(x) = x \log x \qquad x > 0$

$f'(x) = \log x + x \cdot \frac{1}{x} = \log x + 1$

$f''(x) = \frac{1}{x} > 0$

$\therefore$ $x \log x$ is strictly convex function

Ex 4: $f(x) = \log x \qquad x > 0$

$f'(x) = \frac{1}{x} \qquad f''(x) = -\frac{1}{x^2} < 0$

$\therefore$ $-\log x$ is strictly convex

or

$\log x$ is strictly concave.

Ex 5: $f(x) = \sqrt{x} \qquad x > 0$

$f'(x) = \frac{1}{2\sqrt{x}} = \frac{1}{2} x^{-1/2} \qquad f''(x) = -\frac{1}{4} \cdot \frac{1}{(\sqrt{x})^3} = -\frac{1}{4} \frac{1}{x\sqrt{x}} < 0$

$\therefore$ $\sqrt{x}$ is strictly concave.

Defn: The entropy $H(X)$ of a discrete random variable X is defined by

$$H(X) = -\sum_{x \in X} p(x) \log p(x)$$ concave.

$$0 \log 0 = 0$$

Extend this to n-dimension case.

If Hessian of a function is positive semidefinite (positive definite) over a convex set $C \subset R^n$, then function $F(x)$ over C is convex (or strictly convex).

Proof:

$$F(x) = F(x_0) + \nabla F^T(x_0)(x - x_0) + \underbrace{\tfrac{1}{2}(x-x_0)^T H(\bar{x})(x - x_0)}_{\geq 0}$$

$\bar{x} \in [x_0, x]$ $\in C$

$x_0 \in C$

$x \in C$

$$\boxed{F(x) \geq F(x_0) + \nabla F^T(x_0)(x - x_0)}$$

Choose $x_1 \in C$, $x_2 \in C$

define $x_0 = \lambda x_1 + \overline{1-\lambda}\, x_2$ $\lambda \in [0,1]$

Choose one $x = x_1$; next $x = x_2$.

Calculate $(x - x_0)$ & substitute in the inequality

Then multiply first resulting inequality by λ and second one by $(1-\lambda)$ and add.

$$F(\lambda x_1 + \overline{1-\lambda}\, x_2) \leq \lambda F(x_1) + (1-\lambda) F(x_2)$$

$\lambda \in [0,1]$

Hence $F(x)$ is convex.