- Convex Functions
  - $x \in R$
    - $ax+b$
    - e^x
    - $x \log x$, $x>0$
  - $x \in R^n$
    - affine: $f(x) = a^T x + b$
    - norm: $\|a\|_p + \|b\|_p \geq \|a+b\|_p$
    - spectral: $f(x) = \|X\|_2 \equiv \sigma_{max}(X) \, [\underbrace{\lambda_{max}}_{\text{eigenvalue}}(X^T X)]^{\frac{1}{2}}$

- Concave Function

  $f: S_{++}^n \to R$

  $f(x) = \log(\det(x))$

  <proof> Let $g(t) = f(x + t \cdot v)$, $\forall x \in S_{++}^n$, $\forall v \in S^n$, $0 \le t \le \varepsilon$

  $$g(t) = \log(\det(x + tv))$$
  $$= \log\left(\det\left(X^{\frac{1}{2}}\left(I + tX^{-\frac{1}{2}} \cdot v \cdot X^{-\frac{1}{2}}\right) X^{\frac{1}{2}}\right)\right)$$
  $$= \log(\det(X)) + \log(\det(I + tX^{-\frac{1}{2}} v X^{-\frac{1}{2}}))$$
  $$= \underbrace{\log(\det(X))}_{\log} + \sum_{i=1}^{n} \underbrace{\log(1 + t \cdot \lambda I)}_{\log}$$

  $\Rightarrow g(t)$ is concave

  $\Rightarrow f(t)$ is concave

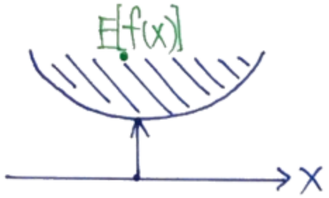
- Max Function

  $f(x) = \max(x_1, \dots, x_n) = \|x\|_\infty$

  <proof>
  $$f(\theta x + (1-\theta) y) = \max_i \theta x_i + (1-\theta) \cdot y_i$$
  $$\le \max_i \theta x_i + \max_i (1-\theta) \cdot y_i$$
  $$= \theta \cdot \max_i x_i + (1-\theta) \cdot \max_i y_i$$
  $$= \theta \cdot f(x) + (1-\theta) \cdot f(y)$$

- Expectation

If $f(x)$ is convex with $p(x)$ as a probability at X. $\int p(x)\,dx = 1$



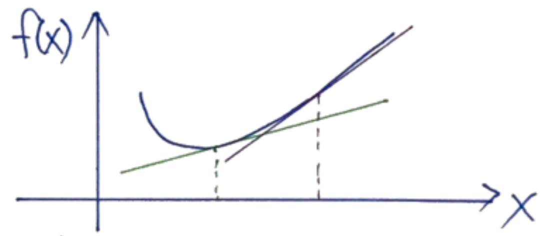
$$f(E[X]) \le E[f(X)]$$

$$\Rightarrow \left[\begin{array}{l} E[X] \equiv \int x \cdot p(x) \cdot dx \\ E[f(x)] \equiv \int f(x) \cdot p(x) \cdot dx \end{array} \right.$$

- First Order Condition

f is differentiable if ① dom f is open, ② $\nabla f(x) = \left(\frac{\partial f(x)}{\partial x_1}, \dots, \frac{\partial f(x)}{\partial x_n}\right)^T$ exists at each $x \in \text{dom } f$

$\Rightarrow f(y) \ge f(x) + \nabla f(x)^T (y - x)$



<proof>

① "$\Rightarrow$": If f is convex,

$$(1-t) f(x) + t \cdot f(y) \ge f((1-t) \cdot x + t \cdot y) \quad \forall\, 0 \le t \le 1$$

$$\Rightarrow -t \cdot f(x) + t \cdot f(y) \ge f(x + t(y-x)) - f(x)$$

$$\Rightarrow t \cdot [f(y) - f(x)] \ge f(x + t(y-x)) - f(x)$$

$$\Rightarrow f(y) - f(x) \ge \frac{f(x + t(y-x)) - f(x)}{t} \xlongequal[t \to 0]{} \nabla f(x)^T \cdot (y - x)$$

② "$\Leftarrow$": Given $f(y) \ge f(x) + \nabla f(x)^T \cdot (y-x)$, $\forall x, y$

Let $z = (1-t)x + ty$, $0 \le t \le 1$

$$f(x) \ge f(z) + \nabla f(z)^T \cdot (x - z)$$

$$f(y) \ge f(z) + \nabla f(z)^T \cdot (y - z)$$

$$(1-t) f(x) + t f(y) \ge (1-t) \cdot f(z) + (1-t) \cdot \nabla f(z)^T \cdot (x - z) + t \cdot f(z) + t \cdot \nabla f(z)^T \cdot (y - z)$$

$$= f(z) + \nabla f(z)^T \cdot [(1-t) \cdot x + ty - (1-t) \cdot z - tz]$$

$$= f(z) = f((1-t) \cdot x + t \cdot y)$$

- Second Order Condition

$$f(y) = f(x) + \nabla f(x)^T (y-x) + \frac{1}{2} \cdot (y-x)^T \cdot \nabla^2 f(z) \cdot (y-x)$$

$\exists t,\ 0 \le t \le 1$, s.t. $z = (1-t)\cdot x + t \cdot y$

$\underbrace{\nabla f^2(x) = [f_{ij}(x)]}_{\text{Lagrange Remainder}}$, where $f_{ij}(x) = \dfrac{\partial^2 f(x)}{\partial x_i \partial x_j}$

$\Rightarrow f(x)$ is convex iff $\nabla^2 f(x) \ge 0$, $\forall x \in \text{dom} f$

- Example
  - Quadratic Function

    $f(x) = \frac{1}{2} x^T P x + q^T x + r$, $P \in S^n$

    $\nabla f(x) = Px + q$

    $\nabla^2 f(x) = P$ * If f is convex, $\nabla^2 f(x) = P \ge 0$. Thus, P is semi-definite

  - Least Square

    $f(x) = \|Ax - b\|_2^2 = (Ax-b)^T (Ax-b)$

    $\nabla f(x) = 2A^T A x - 2A^T b$

    $\nabla^2 f(x) = 2A^T A$ * No matter what A and b are, $A^T A$ is always symmetric and semi-definite

  - Entropy

    $f(x) = x \log x$

    $f'(x) = \log x + 1$

    $f''(x) = \frac{1}{x}$ * dom f : $x > 0 \Rightarrow f''(x) > 0 \Rightarrow f(x)$ is convex

  - Quadratic-over-linear

    $f(x, y) = \frac{x^2}{y}$, $y > 0$

    $\nabla f(x, y) = \left(\frac{2x}{y}, \frac{-x^2}{y^2}\right)$

    $\nabla^2 f(x, y) = \begin{bmatrix} \frac{2}{y} & \frac{-2x}{y^2} \\ \frac{-2x}{y^2} & \frac{2x^2}{y^3} \end{bmatrix}$ $\left[\begin{array}{l} ① \ \frac{2}{y} > 0, \ \frac{2x^2}{y^3} > 0 \\ ② \ \frac{2}{y} \cdot \frac{2x^2}{y^3} - \left(\frac{-2x}{y^2}\right) \cdot \left(\frac{-2x}{y^2}\right) = 0 \end{array}\right.$ * $f(x, y)$ is convex

- log-sum-op (softmax)

$$f(x) = \log \sum_{k=1}^{n} e^{x_k}$$

$$\frac{\partial f(x)}{\partial x_i} = \frac{e^{x_i}}{\sum_{k=1}^{n} e^{x_k}}$$

$$\frac{\partial^2 f(x)}{\partial x_i^2} = \frac{e^{x_i} \cdot \left(\sum_{k=1}^{n} e^{x_k}\right) - e^{x_i} \cdot e^{x_i}}{\left(\sum_{k=1}^{n} e^{x_k}\right)^2}$$

$$\frac{\partial f(x)}{\partial x_i x_j} = \frac{0 \cdot \left(\sum_{k=1}^{n} e^{x_k}\right) - e^{x_i} \cdot e^{x_j}}{\left(\sum_{k=1}^{n} e^{x_k}\right)^2}$$

Assume $z_k = e^{x_k}$,

$$\Rightarrow \nabla^2 f(x) = \frac{1}{1^T z}\left[\mathrm{diag}(z) - z z^T\right]$$

$$\Rightarrow v^T \cdot \nabla^2 f(x) \cdot v = \frac{\left(\sum_k z_k v_k^2\right)\left(\sum_k z_k\right) - \left(\sum_k v_k z_k\right)^2}{\left(\sum_k z_k\right)^2} \geq 0, \ \forall v$$

(from Cauchy-Schwarz, $\sum_k z_k v_k^2 \cdot \sum_k z_k \geq \left(\sum_k v_k z_k\right)^2$)