- Conjugate Function

$f(x) = \log \sum_{i=1}^{n} e^{x_i}$

$f^*(y) = \sup_x \left(y^T x - \log \sum_{i=1}^{n} e^{x_i} \right)$

let $g(x) = y^T x - \log \sum_{i=1}^{n} e^{x_i}$

$\frac{\partial g(x)}{\partial x_i} = y_i - \frac{e^{x_i}}{\sum_{i=1}^{n} e^{x_i}} = 0 \Rightarrow y_i = \frac{e^{x_i}}{\sum_{i=1}^{n} e^{x_i}} \Rightarrow x_i = \log y_i + \log\left(\sum_{i=1}^{n} e^{x_i}\right)$

Thus, $f^*(y) = \sum_{i=1}^{n} y_i \cdot \left[\log y_i + \log \sum_{i=1}^{n} e^{x_i}\right] - \log \sum_{i=1}^{n} e^{x_i}$

$= \sum_{i=1}^{n} y_i \log y_i + 1 \cdot \log \sum_{i=1}^{n} e^{x_i} - \log \sum_{i=1}^{n} e^{x_i}$ (since $\sum_{i=1}^{n} y_i = 1$),

if (1) $y_i \geq 0$, or $f^*(y) \to \infty$

(2) $\sum_{i=1}^{n} y_i = 0$, or $f^*(y) \to \infty$

- Formulation:

$\min f_0(x)$,

subject to $f_i(x) = 0$, $i = 1, \ldots, m$
$h_i(x) = 0$, $i = 1, \ldots, p$,

$x \in R^n$
$\begin{cases} f_0 : R^n \to R \\ f_i : R^n \to R \\ h_i : R^n \to R \end{cases}$, $f_0, \ldots, f_m$ are convex, $D : D_{f_0} \cap D_{f_1} \cap \cdots \cap D_{f_n}$

- (I) Feasible set of a convex optimization problem is convex
- (II) Local optimal solution is global optimal

* (1) Local optimal

for a x, $\exists r > 0$, s.t. $\|z - x\|_2 \leq r$, $f_0(z) \geq f_0(x)$, $\forall z \in D$

* (2) local optimal $\Rightarrow$ global optimal

<proof: by contradiction>

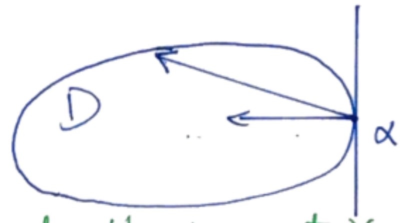
Suppose $\exists y \in D$ s.t. $f_0(y) < f_0(x)$

$f_0(x) > \theta f_0(y) + (1-\theta) \cdot f_0(x)$, $\theta > 0$

$\geq f_0(\theta y + (1-\theta) \cdot x)$

As $\theta \to 0$, $\theta y + (1-\theta) x$ falls into the ball $\|z - x\|_2 \leq r$, $\theta y + (1-\theta) x \in D$

- (III) If $\nabla f_0^T(x)(y-x) \geq 0, \ \forall x, y \in D$,
  Then x is an optimal solution.



  * $(y-x)^T \nabla f_0(x)$ defines a supporting hyperplane to feasible solution at x

- (IV) Unconstrainted Problem
  If $\nabla f_0(x) = 0$, $x \in \operatorname{dom} f$, then x is an optimal solution.

- (V) Equality Constraint Problem

$$\begin{cases} \min f_0(x) \\ \text{subject to } Ax = b \end{cases}$$

  x is an optimal solution iff $x \in \operatorname{dom} f_0$, $Ax=b$, $\nabla f_0(x) + \underbrace{A^T\nu}_{\text{dual cone}} = 0$, $\exists \nu \in R^+$

Example:

$\min f(x) = x_1^2 + x_2^2$

s.t. $Ax = b : [2 \ \ 1]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 3$

$$\Rightarrow \nabla f_0(x) = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix} \Rightarrow \nabla f_0(x) + A^T\nu = 0 \Rightarrow \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \end{bmatrix}\nu = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 2 & 0 & 2 \\ 0 & 2 & 1 \\ 2 & 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \nu \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix} \Rightarrow \begin{bmatrix} x_1 \\ x_2 \\ \nu \end{bmatrix} = \begin{bmatrix} \frac{6}{5} \\ \frac{3}{5} \\ -\frac{6}{5} \end{bmatrix}$$

- Use (III) to prove (V)

  x is optimal, $\nabla f_0(x)^T(y-x) \geq 0$, $\forall y \in D$, i.e. $y = x + w$, s.t. $Aw = 0$

  Thus, $\begin{array}{ll} \nabla f_0(x)^T w \geq 0, & \forall Aw = 0 \\ -\nabla f_0(x)^T w \geq 0, & \forall A(-w) = 0 \end{array} \Rightarrow$

  $\Rightarrow \nabla f_0(x)^T w = 0$, $\forall Aw = 0$ implies $\nabla f_0(x) + A^T\nu = 0$