Convex Optimization - Homework 1:
Let $f : \mathbb{R}^n \to \mathbb{R}$, $\text{dom}(f) = \mathbb{R}^n_{++} \cap \{x : x \in \sum_{i=1}^n e^{-x_i} \leq 1\}$ be

$$f(x) = \prod_{i=1}^n (1 - e^{-x_i}).$$

Prove that this function is concave.

Solution.

We will start by showing that $\text{dom}(f)$ is convex. Let's take points x and y which satisfy the given inequality. We will show that for $\theta \in (0,1)$, the point $\theta x + (1-\theta)y$ also satisfies the given inequality. From Hölder's inequality [for $(r,s) \in \mathbb{R}_+$: $(\sum_{k=1}^n |x_k|^r |y_k|^s)^{r+s} \leq (\sum_{k=1}^n |x_k|^{r+s})^r (\sum_{k=1}^n |y_k|^{r+s})^s$], we know that:

$$\sum_{i=1}^n \frac{1}{e^{\theta x_i + (1-\theta)y_i}} = \left(\sum_{i=1}^n \left|\frac{1}{e^{x_i}}\right|^\theta \cdot \left|\frac{1}{e^{y_i}}\right|^{1-\theta}\right)^1 \leq \left(\sum_{i=1}^n \left|\frac{1}{e^{x_i}}\right|\right)^\theta \cdot \left(\sum_{i=1}^n \left|\frac{1}{e^{y_i}}\right|\right)^{1-\theta} \leq 1 \cdot 1 = 1$$

So our point satisfies the given inequality. Therefore, the set of points which satisfy the given inequality is convex. Also, the set $\mathbb{R}^n_{++}$ is convex. Thus, our $\text{dom}(f)$ is convex because the intersection of convex sets is convex.

Now we are computing the hessian of function $g = -f$. After simple calculations, we get:

$$\frac{\partial^2 g}{\partial x_i^2} = e^{-x_i} \cdot \prod_{j \neq i} (1 - e^{-x_j}) \qquad \frac{\partial^2 g}{\partial x_i x_j} = -e^{-x_i - x_j} \cdot \prod_{k \neq i,j} (1 - e^{-x_k})$$

We will show that $\text{Hess}(g) = (v \cdot v^T) \circ (\text{diag}(e^{x_i}) - J_n)$ for such a v that $v_i = e^{-i}\sqrt{\frac{\prod_{j \neq i}(1-e^{-x_j})}{(1-e^{-x_i})}}$.

$$v_i^2 \cdot (e^{x_i} - 1) = e^{-2x_i} \frac{\prod_{j \neq i}(1 - e^{-x_j})}{(1 - e^{-x_i})} \cdot \frac{(1 - e^{-x_i})}{e^{-x_i}} = e^{-x_i} \cdot \prod_{j \neq i}(1 - e^{-x_j}) = \frac{\partial^2 g}{\partial x_i^2}$$

$$v_i \cdot v_j \cdot (-1) = -e^{-x_i}\sqrt{\frac{\prod_{l \neq i}(1 - e^{-x_l})}{(1 - e^{-x_i})}} \cdot e^{-x_j}\sqrt{\frac{\prod_{k \neq j}(1 - e^{-x_k})}{(1 - e^{-x_j})}} = -e^{x_i - x_j} \prod_{m \neq i,j}(1 - e^{-x_m}) = \frac{\partial^2 g}{\partial x_i x_j}$$

So the matrix $\text{Hess}(g)$ actually has this form.

Now we want to show that vv^T and $(\text{diag}(e^{x_i}) - J_n)$ are PSD matrices. We can see that $x^T(vv^T)x = (v^Tx)^T \cdot (v^Tx) \geq 0$.

For the second matrix, we have:

$$a^T(\text{diag}(e^{x_i}) - J_n)a = \sum_i (e^{x_i} - 1)a_i^2 - \sum_{i \neq j} a_i a_j = \sum_i e^{x_i} a_i^2 - (\sum_i a_i)^2$$

From Cauchy–Schwarz inequality, we have:

$$\left(\sum_i e^{x_i} a_i^2\right)\left(\sum_i e^{-x_i}\right) \geq \left(\sum_i e^{x_i/2} a_i \cdot e^{-x_i/2}\right)^2 = \left(\sum_i a_i\right)^2$$

Therefore, $\sum_i e^{x_i} a_i^2 - (\sum_i a_i)^2 \geq \sum_i e^{x_i} a_i^2 - (\sum_i e^{x_i} a_i^2)(\sum_i e^{-x_i}) = (\sum_i e^{x_i} a_i^2)(1 - \sum_i e^{-x_i}) \geq 0$. The last inequality is implied by our assumption about the domain.

Finally, we have proved that the matrix $(\mathrm{diag}(e^{x_i}) - J_n)$ is PSD. Therefore, $\mathrm{Hess}(g)$ is PSD as a coordinate-wise matrix product of PSD matrices. This implies that g is a convex function, so f is concave. $\square$