Venkateswar Reddy M.
CTO, Brillium Technologies
"...dare to dream; care to win..."

Date: **28-Apr-24**

EE908 Assignment-4 Solution

---

eMasters in Communication Systems, IITK
EE908: Optimization in SPCOM
Instructor: Prof. Ketan Rajawat
Student Name: Venkateswar Reddy Melachervu
Roll No: 23156022

Q1. Using the concavity of the logarithm, show that $x^\theta y^{1-\theta} \leq \theta x + (1-\theta)y$.
Solution:
Given inequality: $x^\theta y^{1-\theta} \leq \theta x + (1-\theta)y$
Taking log both sides:
$\log\left(x^\theta y^{1-\theta}\right) \leq \log(\theta x + (1-\theta)y)$
$\Rightarrow \boldsymbol{\theta\, log(x) + (1-\theta)\, log\, y \leq log(\theta x + (1-\theta)y)}$

Now let's make use of log concavity and Jensen's inequality to prove the above inequality.

∴Jensen's inequality for concavity of a function,
$f(\Sigma_{i=1}^n \theta_i x_i) \geq \Sigma_{i=1}^n \theta_i f(x_i)\, , \Sigma_{i=1}^n \theta_i = 1, \forall x_i \in dom\ f$

Log(x) is a concave function and for any $x, y \in dom\ (\log), \theta \in [0,1]$
Applying Jensen's inequality:
$log(\theta x + (1-\theta)y) \geq \theta\, log(x) + (1-\theta)\log(y)$
$\therefore \boldsymbol{\theta\, log(x) + (1-\theta)\log(y) \leq log(\theta x + (1-\theta)y)}$
QED

Q2. Show that the harmonic mean $f(x) = \left(\Sigma_{i=1}^n \frac{1}{x_i}\right)^{-1}$ is concave.

Solution:
Let's use the second order conditions $\nabla f''(x) < 0$ (opposite of PSD) to prove $f(x)$ is concave

$$\frac{\partial f(x)}{\partial x_i} = (\Sigma_{i=1}^n x_i^{-1})^{-2} x_i^{-2} = \left(\frac{f(x)}{x_i}\right)^2$$

$$\frac{\partial^2 f(x)}{\partial x_i \partial x_j} = \frac{2}{x_i}\left(\frac{f(x)}{x_i}\right)\left(\frac{f(x)}{x_j}\right)^2 = \frac{2}{f(x)}\left(\frac{f(x)^2}{x_i x_j}\right)^2 \text{ for } i \neq j$$

$$\frac{\partial^2 f(x)}{\partial x_i^2} = \frac{2}{f(x)}\left(\frac{f(x)^2}{x_i^2}\right)^2 - \frac{2}{x_i}\left(\frac{f(x)}{x_i}\right)^2 \text{ for } i = j$$

∴ We need to show that

$$y^T \nabla^2 f(x) y = \frac{2}{f(x)}\left(\left(\Sigma_{i=1}^n \frac{y_i f(x)^2}{x_i^2}\right)^2 - \left(\Sigma_{i=1}^n \frac{y_i^2 f(x)^3}{x_i^3}\right)\right) \leq 0 \text{ for concavity of } f(x)$$

The above holds true following the Cauchy-Schwartz Inequality $\boldsymbol{a_i^T b_i \leq \|a_i\|_2 \|b_i\|_2}$ where:
$\boldsymbol{a_i = \left(\frac{f(x)}{x_i}\right)^{1/2}, b_i = y_i\left(\frac{f(x)}{x_i}\right)^{3/2}}$ **and $\boldsymbol{\Sigma_i a_i^2 = 1}$**
QED

"...dare to dream; care to win..."

Q3. Prove the reverse Jensen's inequality for a convex f with dom $f = \mathbb{R}^n$, $\lambda_i > 0$ and $\lambda_1 - \Sigma_{i=2}^{n}\lambda_i = 1$
$f(\lambda_1 x_1 - \lambda_2 x_2 - \cdots - \lambda_n x_n) \geq \lambda_1 f(x_1) - \lambda_2 f(x_2) - \cdots - \lambda_n(x_n)$

Solution:

Given the function f is convex function, Jensen's inequality of this function f is:
$\Rightarrow f(\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_n x_n) \leq \lambda_1 f(x_1) + \lambda_2 f(x_2) + \cdots + \lambda_n f(x_n)$
Negating both sides:
$\Rightarrow -f(\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_n x_n) \geq -\lambda_1 f(x_1) - \lambda_2 f(x_2) - \cdots - \lambda_n f(x_n)$
$$\Rightarrow -f\left(\sum_{i=1}^{n} \lambda_i x_i\right) \geq -\sum_{i=1}^{n} \lambda_i f(x_i)$$
Since $\lambda_1 - \Sigma_{i=2}^{n}\lambda_i = 1$
$\Rightarrow -f(\lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_n x_n) \geq -\lambda_1 f(x_1) - (\lambda_1 - 1)$
$-f(\lambda_1 x_1 + \lambda_1 - 1) \geq -\lambda_1 f(x_1) - (\lambda_1 - 1)$
Reversing the signs both sides,
$\boldsymbol{f(\lambda_1 x_1 + \lambda_1 - 1) \leq \lambda_1 f(x_1) + (\lambda_1 - 1)}$
This is valid and holds as per Jensen's inequality
QED

Q4. Give an example of a function $f(x)$ whose epigraph is (a) half-space, (b) norm cone, and (c) polyhedron.

Solution:

Let's define a function $f(x) = \|x\|_2 = \sqrt{x_1^2 + x_2^2}, x \in \mathbb{R}^2$ – Euclidean norm.
This function satisfies all three cases:

(a) Half-space:
- The epigraph of $f(x)$ as a half-space can be represented as $\left\{(x_1, x_2, t) \mid \sqrt{x_1^2 + x_2^2} \geq t\right\}$
- This corresponds to the region above or on the circle centered at the origin with radius t

(b) Norm cone
- The epigraph of $f(x)$ as a norm cone can be represented as $\left\{(x_1, x_2, t) \mid \sqrt{x_1^2 + x_2^2} \leq t\right\}$
- This represents the cone-like region expanding from the origin outward.

(c) Polyhedron
- The epigraph of $f(x)$ as a polyhedron is the set of points (x_1, x_2, t) such that $\sqrt{x_1^2 + x_2^2} \leq t$
- This represents circular region in 2D space

So, the function $(x) = \|x\|_2 = \sqrt{x_1^2 + x_2^2}$ satisfies all three cases: it represents a half-space, a norm cone, and a polyhedron.
QED

Q5. Let $x, y \in \mathbb{R}_{++}^n$ be two vectors. We need to show that the Itakura-Saito distance, defined as
$$D_{IS}(x, y) = \sum_{i=1}^{n} \left(\frac{x_i}{y_i} - \log\left(\frac{x_i}{y_i}\right) - 1\right)$$

is always positive, using the following steps:
(a) Show that for a convex differentiable function f, the Bregman divergence:
$$D(x, y) = f(x) - f(y) - \Delta f(y)^T (x - y)$$

is always non-negative.

Solution:

Let's prove that $D(x,y) = f(x) - f(y) - \Delta f(y)^T(x-y) \geq 0$ for all x, y. f is convex differentiable:
So, as per the first order convex condition (gradient-based)
$f(x) \geq f(y) + \nabla f(y)^T(x-y)$
Let's substitute the above equation into Bregman divergence:
$\boldsymbol{D(x,y) = f(x) - f(y) - \nabla f(y)^T(x-y)}$
$\boldsymbol{\geq [f(y) - \nabla f(y)^T(x-y)] - f(y) - \nabla f(y)^T(x-y)}$
$\boldsymbol{\geq 0}$
QED

(b) Show that for the convex function $f(x) = -\Sigma_{i=1}^n \log(x_i)$, it holds that $D(x,y) = D_{IS}(x,y)$.

Solution:

Given $f(x) = -\Sigma_{i=1}^n \log(x_i)$,
$D(x,y) = f(x) - f(y) - \nabla f(y)^T(x-y)$
$\therefore \nabla f(x) = \left(-\frac{1}{x_1}, -\frac{1}{x_2}, \ldots -\frac{1}{x_n}\right)$
$\therefore D(x,y) = -\Sigma_{i=1}^n \log(x_i) + \Sigma_{i=1}^n \log(y_i) - \left(-\frac{1}{x_1}, -\frac{1}{x_2}, \ldots -\frac{1}{x_n}\right)^T (x-y)$
$= -\Sigma_{i=1}^n \log(x_i) + \Sigma_{i=1}^n \log(y_i) + \Sigma_{i=1}^n \dfrac{x_i}{y_i} - n$
Let's compute $D_{IS}(x,y) = \Sigma_{i=1}^n \left(\frac{x_i}{y_i} - \log\left(\frac{x_i}{y_i}\right) - 1\right)$
$$= \sum_{i=1}^n \left(\frac{x_i}{y_i} - \log\left(\frac{x_i}{y_i}\right)\right) - n$$
Let's compare both:
$D(x,y) = -\Sigma_{i=1}^n \log(x_i) + \Sigma_{i=1}^n \log(y_i) + \Sigma_{i=1}^n \dfrac{x_i}{y_i} - n$
$= -(\Sigma_{i=1}^n \log(x_i) - \Sigma_{i=1}^n \log(y_i)) + \Sigma_{i=1}^n \dfrac{x_i}{y_i} - n$
$= \Sigma_{i=1}^n \dfrac{x_i}{y_i} - (\Sigma_{i=1}^n \log(x_i) - \Sigma_{i=1}^n \log(y_i)) - n$
$= \Sigma_{i=1}^n \dfrac{x_i}{y_i} - \left(\Sigma_{i=1}^n \log\left(\dfrac{x_i}{yi}\right)\right) - n$
$$\boldsymbol{= \sum_{i=1}^n \left(\frac{x_i}{y_i} - \log\left(\frac{x_i}{yi}\right)\right) - n = D_{IS}(x,y)}$$
$\boldsymbol{\therefore D(x,y) = D_{IS}(x,y)}$
QED

(c) Along similar lines, prove that the generalized KL divergence:
$$D_{KL}(x,y) = \Sigma_{i=1}^n \left(x_i \log\left(\frac{x_i}{y_i}\right) - x_i + y_i\right)$$
is always positive.

Solution:

We aim to prove that $D_{KL}(x,y) \geq 0$ for all non-negative x, y i.e.
$$D_{KL}(x,y) = \sum_{i=1}^n \left(x_i \log\left(\frac{x_i}{y_i}\right) - x_i + y_i\right) \geq 0$$
Jensen's inequality for a convex function $f(x) = -\log(x)$

C-501, Salarpuria Serenity, 5th Main, Sector 7, HSR Layout, Bengaluru 560102 KA India
Mobile: +91 97012 22130, Email: vmelachervu@gmail.com, vmela23@iitk.ac.in, Website: www.linkedin.com/in/vmelachervu

$$-\log\left(\sum_{i=1}^{n}\theta_i\frac{x_i}{y_i}\right)\le\sum_{i=1}^{n}\theta_i\left(-\log\left(\frac{x_i}{y_i}\right)\right)$$

Rewriting the above:

$$-\log\left(\sum_{i=1}^{n}\theta_i\frac{x_i}{y_i}\right)\le\sum_{i=1}^{n}\theta_i(-\log(x_i)+\log(y_i))$$

$$\Rightarrow -\log\left(\sum_{i=1}^{n}\theta_i\frac{x_i}{y_i}\right)\le-\sum_{i=1}^{n}\theta_i\log(x_i)+\sum_{i=1}^{n}\theta_i\log(y_i)$$

$$\Rightarrow \log\left(\sum_{i=1}^{n}\theta_i\frac{x_i}{y_i}\right)\ge\sum_{i=1}^{n}\theta_i\log(x_i)-\sum_{i=1}^{n}\theta_i\log(y_i)$$

Substituting the inequality into KL divergence:

$$D_{KL}(x,y)=\sum_{i=1}^{n}\left(x_i\log\left(\frac{x_i}{y_i}\right)-x_i+y_i\right)\ge\sum_{i=1}^{n}x_i\log(x_i)+\sum_{i=1}^{n}x_i\log(y_i)-\sum_{i=1}^{n}x_i+\sum_{i=1}^{n}y_i$$

Simplifying:

$$\boldsymbol{D_{KL}(x,y)\ge\sum_{i=1}^{n}y_i-\sum_{i=1}^{n}x_i}$$

Since $\boldsymbol{x_i, y_i}$ are non-negative for all $\boldsymbol{i}$, $\boldsymbol{y_i - x_i}$ is non-negative
QED

--- End of the Document---

C-501, Salarpuria Serenity, 5th Main, Sector 7, HSR Layout, Bengaluru 560102 KA India
Mobile: +91 97012 22130, Email: vmelachervu@gmail.com, vmela23@iitk.ac.in, Website: www.linkedin.com/in/vmelachervu