Convex Optimization - AI2101
Assignment - VI

Bhuvan Chandra K
AI23BTECH11013

April 27, 2025

Question 3.36 (a)

Derive the conjugates of the following functions.
(a) Max function: $f(x) = \max_{i=1,\dots,n} x_i$ on $\mathbb{R}^n$.

Solution:

Let

$$f(x) = \max_{i=1,2,\dots,n} x_i \quad \text{for } x \in \mathbb{R}^n$$

We need to find the convex conjugate $f^*(y)$ defined as:

$$f^*(y) = \sup_{x \in \mathbb{R}^n} (y^T x - f(x))$$

$$f^*(y) = \sup_{x \in \mathbb{R}^n} \left(y^T x - \max_i x_i \right)$$

Here we start with a constant vector $x = t.\mathbf{1}$.
When we do this, the supremum becomes a one dimensional problem:

- All components of x are equal: $x_1 = x_2 = \cdots = x_n = t$
- $\max_i x_i = t$

Then:

$$y^T x = t \sum_{i=1}^{n} y_i = t \cdot \mathbf{1}^T y$$

i.e.,

$$y^T x - \max_i x_i = t(\mathbf{1}^T y - 1)$$

$$f^*(y) \geq \sup_{t \in \mathbb{R}} t(\mathbf{1}^T y - 1)$$

We use this lower bound from constant vector.

If $\mathbf{1}^T y > 1$, then $t(\mathbf{1}^T y - 1) \to +\infty$ as $t \to +\infty$
If $\mathbf{1}^T y < 1$, then $t(\mathbf{1}^T y - 1) \to +\infty$ as $t \to -\infty$

Both of these cases lead to supremum being $+\infty$, so constraints are not taken.
We can say: if $\mathbf{1}^T y \neq 1$, then $f^*(y) = +\infty$
When $\mathbf{1}^T y = 1$, which means $\sum_{i=1}^{n} y_i = 1$, we get:

$$f^*(y) = \sup_{x \in \mathbb{R}^n} \left(\sum_{i=1}^{n} y_i x_i - \max_i x_i \right)$$

Let's take $M = \max_i x_i \quad (x_i \leq M \quad \forall i)$
Then:

$$\sum_{i=1}^{n} y_i x_i \leq \sum_{i=1}^{n} y_i M$$

$$\sum_{i=1}^{n} y_i x_i \leq M \sum_{i=1}^{n} y_i$$

$$\sum_{i=1}^{n} y_i x_i \leq M$$

Subtracting $\max_i x_i$ on both sides of the inequality:

$$\sum_{i=1}^{n} y_i x_i - \max_i x_i \leq M - M$$

$$\sum_{i=1}^{n} y_i x_i - \max_i x_i \leq 0$$

So,

$$f^*(y) \leq 0$$

We got this from the original expression of $f^*(y)$.
Now using $\mathbf{1}^T y = 1$ in the first inequality:

$$f^*(y) \geq \sup_{t \in \mathbb{R}} \Big(t(\mathbf{1}^T y - 1) \Big)$$

$$f^*(y) \geq \sup_{t \in \mathbb{R}} (t(0))$$

$$f^*(y) \geq 0$$

We have $f^*(y) \leq 0$ and $f^*(y) \geq 0 \Rightarrow f^*(y) = 0$
But there is one other constraint we have to restrict. Let $y_j < 0$ for some j.
In this case, let $x_j \to -\infty$ and rest of x_k $(k \neq j)$ are bounded and their maximum is m.
This gives:

$$y^T x \to +\infty, \quad \max_i x_i = m$$

$$f^*(y) \to +\infty - m \to +\infty$$

So $y_j < 0$ is not in the domain of f^*.

$$\therefore f^*(y) = \begin{cases} 0 & \text{if } y \succeq 0,\ \mathbf{1}^T y = 1 \\ \infty & \text{otherwise} \end{cases}$$

Note: Even if x is not constant, $y^T x$, which is the weighted average of x components with weights y_i, and since $\mathbf{1}^T y = 1$,

$$y^T x \leq \max_i x_i \quad (\text{for } y \geq 0, \mathbf{1}^T y = 1)$$

So,

$$y^T x - \max_i x_i \leq 0$$

We still get supremum over all such $x \in \mathbb{R}^n$ as o when $y \succeq 0$ and $\mathbf{1}^T y = 1$.

Question 3.36 (b)

Derive the conjugates of the following functions.
(b) Sum of largest elements: $f(x) = \sum_{i=1}^{r} x_{[i]}$ on $\mathbb{R}^n$.

Solution:

$$f(x) = \sum_{i=1}^{r} x_{[i]} \quad \text{on } \mathbb{R}^n$$

where

$$x_{[1]} \geq x_{[2]} \geq x_{[3]} \geq \cdots \geq x_{[n-1]} \geq x_{[n]}$$

We need the convex conjugate $f^*(y)$ defined as

$$f^*(y) = \sup_{x \in \mathbb{R}^n} \left(y^T x - f(x) \right)$$

$$f^*(y) = \sup_{x \in \mathbb{R}^n} \left(y^T x - \sum_{i=1}^{r} x_{[i]} \right)$$

Define:

$$g(x) = y^T x - \sum_{i=1}^{r} x_{[i]}$$

Let $y_j > 1$ for some j. Also $x_j = k > 0$ and all other components of x are zero.
If $k > 0$, then $x_{[1]} = k,\ x_{[2]} = 0, \ldots, x_{[r]} = 0$,

$$\sum_{i=1}^{r} x_{[i]} = k$$

$$g(x) = k \cdot y_j - k = k(y_j - 1)$$

If $k \to +\infty \Rightarrow g(x) \to +\infty$, if $y_j > 1$
i.e. $y_j > 1$ is not in the domain of $f^*(y)$

$$y_j \leq 1 \quad \forall j$$

Let $y_j < 0$ for some j. Also $x_j = k$ and all other components of x are zero.
If $k < 0$, then $x_{[1]} = 0,\ x_{[2]} = 0,\ \ldots,\ x_{[r]} = k$,

$$\sum_{i=1}^{r} x_{[i]} = k$$

$$g(x) = ky_j - k = -k(1 - y_j)$$

If $-k \to +\infty \Rightarrow g(x) \to +\infty$ if $y_j < 0$
i.e. $y_j < 0$ is also not in the domain of $f^*(y)$

$$y_j \geq 0 \quad \forall j$$

We have $y_j \geq 0\ \forall j$ and $y_j \leq 1\ \forall j \Rightarrow 0 \leq y_j \leq 1 \quad \forall j$
Now let's choose a constant vector $x = k \cdot \mathbf{1}$
With this assumption, this becomes a one dimensional problem.
Then,

$$\sum_{i=1}^{r} x_{[i]} = r \cdot k$$

$$y^T x = k \cdot \sum_{i} y_i = k(\mathbf{1}^T y)$$

$$g(x) = y^T x - \sum_{i=1}^{r} x_{[i]} = k(\mathbf{1}^T y - r)$$

If $\mathbf{1}^T y > r$, then $k(\mathbf{1}^T y - r) \to +\infty$ as $k \to +\infty$
If $\mathbf{1}^T y < r$, then $k(\mathbf{1}^T y - r) \to +\infty$ as $k \to -\infty$

In both these cases, supremum reaches $+\infty$, so these aren't considered.
Hence, the only y in the domain of $f^*(y)$ must satisfy:

$$\mathbf{0} \preceq y \preceq \mathbf{1}, \quad \text{and} \quad \mathbf{1}^T y = r$$

Let's denote, as

$$x_{\pi(1)} = x_{[1]}, \quad x_{\pi(2)} = x_{[2]}, \quad \ldots, \quad x_{\pi(n)} = x_{[n]}$$

$$y^T x = \sum_{j=1}^{n} y_j x_j = \sum_{j=1}^{n} y_{\pi(j)} x_{\pi(j)}$$

$$y^T x = \sum_{j=1}^{n} y_{\pi(j)} x_{[j]}$$

$y^T x$ will be maximum when the largest values of $y_{\pi(j)}$ are paired with those of $x_{[j]}$.
So, top r values of $y_{\pi(j)}$ will be 1 and others zero.

$$\sum_{j=1}^{n} y_{\pi(j)} x_{[j]} \leq \sum_{j=1}^{n} x_{[j]}$$

So,

$$y^T x \leq \sum_{i=1}^{r} x_{[i]} \quad \forall x \in \mathbb{R}^n$$

$$y^T x - \sum_{i=1}^{r} x_{[i]} \leq 0 \quad \forall x \in \mathbb{R}^n$$

$$f^*(y) = \sup_x \left(y^T x - \sum_{i=1}^{r} x_{[i]} \right)$$

So , $f^*(y) = 0$ with the final constraints

$$\therefore \quad f^*(y) = \begin{cases} 0 & \text{if } 0 \preceq y \preceq 1,\ \mathbf{1}^T y = r \\ \infty & \text{otherwise} \end{cases}$$

Question 3.36(d)

Derive the conjugates of the following functions.
(d) Power function: $f(x) = x^p$ on $\mathbb{R}_{++}$, where $p > 1$.

Solution:

$$f(x) = x^p \quad \text{for } x \in \mathbb{R}_{++},\ p > 1$$

We define the conjugate of $f(x)$ as

$$f^*(y) = \sup_{x>0} \left[yx - x^p \right]$$

We want to find the maximum of $yx - x^p$ over $x > 0$. We take the derivative and proceed.
Let ,

$$g(x) = yx - x^p$$

$$g'(x) = \frac{d}{dx}(yx - x^p) = y - px^{p-1}$$

So for the critical point, set $g'(x) = 0$:

$$y - px^{p-1} = 0 \Rightarrow x = \left(\frac{y}{p}\right)^{\frac{1}{p-1}}$$

The value of x that results in maximum of $yx - x^p$ is

$$\hat{x} = \left(\frac{y}{p}\right)^{\frac{1}{p-1}}$$

Substituting $\hat{x}$ in $f^*(y)$,

$$f^*(y) = y \cdot \hat{x} - (\hat{x})^p$$

$$f^*(y) = y\left(\frac{y}{p}\right)^{\frac{1}{p-1}} - \left(\frac{y}{p}\right)^{\frac{p}{p-1}}$$

$$f^*(y) = (p-1)\left(\frac{y}{p}\right)^{\frac{p}{p-1}}$$

All of this was for $y > 0$, but for $y \leq 0$, the supremum of $g(x)$ occurs at $x = 0$, because yx will be negative if $x > 0$, which will result in $yx - x^p$ being negative.
But when $x = 0 \Rightarrow yx - x^p = 0$

$$\therefore \quad f^*(y) = \begin{cases} 0 & \text{if } y \leq 0 \\ (p-1)\left(\frac{y}{p}\right)^{\frac{p}{p-1}} & \text{if } y > 0 \end{cases}$$

Question 3.42

Approximation width. Let $f_0, \ldots, f_n : \mathbb{R} \to \mathbb{R}$ be given continuous functions. We consider the problem of approximating f_0 as a linear combination of $f_1, \ldots, f_n$. For $x \in \mathbb{R}^n$, we say that

$$f = x_1 f_1 + \cdots + x_n f_n$$

approximates f_0 with tolerance $\epsilon > 0$ over the interval $[0, T]$ if

$$|f(t) - f_0(t)| \leq \epsilon \quad \text{for } 0 \leq t \leq T.$$

Now we choose a fixed tolerance $\epsilon > 0$ and define the *approximation width* as the largest T such that f approximates f_0 over the interval $[0, T]$:

$$W(x) = \sup \{T \mid |x_1 f_1(t) + \cdots + x_n f_n(t) - f_0(t)| \leq \epsilon \text{ for } 0 \leq t \leq T\}.$$

Show that W is quasiconcave.

Solution:

Given, $f_0, f_1, \ldots, f_n : \mathbb{R} \to \mathbb{R}$ are continuous functions.
We want to approximate f_0 as a linear combination of $f_1, f_2, \ldots, f_n$.
For a given tolerance $\epsilon > 0$, the approximation width $W(x)$ is defined as

$$W(x) = \sup \{T \mid |x_1 f_1(t) + x_2 f_2(t) + \cdots + x_n f_n(t) - f_0(t)| \leq \epsilon \quad \forall t \in [0, T]\}$$

We need to show that $W(x)$ is quasi-concave.
The function $W(x)$ is quasi-concave if, for any $\alpha \in \mathbb{R}$, the sublevel set

$$S_\alpha = \{x \in \mathbb{R}^n \mid W(x) \geq \alpha\}$$

is convex. So for any two points $x, y \in \mathbb{R}^n$ and any $\alpha \in \mathbb{R}$, S_α is convex.
Thus, for all $z \in S_\alpha$, the linear combination $x_1 f_1(t) + \cdots + x_n f_n(t)$ approximates $f_0(t)$ with tolerance ϵ over an interval of length α.
That is, $\forall x \in S_\alpha$, there exists $T \geq \alpha$ such that

$$|x_1 f_1(t) + \cdots + x_n f_n(t) - f_0(t)| \leq \epsilon \quad \forall t \in [0, T]$$

For S_α to be convex, if $x, y \in S_\alpha$, then for any $\lambda \in [0, 1]$ the point $z = \lambda x + (1 - \lambda) y$ must also be in S_α.
Let $x \in S_\alpha$, $y \in S_\alpha$. There exist intervals $[0, T_1]$, $[0, T_2]$ respectively for x and y with

$$T_1 \geq \alpha \quad \text{and} \quad T_2 \geq \alpha$$

such that for all $t \in [0, T_1]$ and $t \in [0, T_2]$,

$$|x_1 f_1(t) + \cdots + x_n f_n(t) - f_0(t)| \leq \epsilon$$

$$|y_1 f_1(t) + \cdots + y_n f_n(t) - f_0(t)| \leq \epsilon$$

Now, define

$$z = \lambda x + (1 - \lambda) y$$

thus

$$z_1f_1(t)+\cdots+z_nf_n(t)=\lambda(x_1f_1(t)+\cdots+x_nf_n(t))$$
$$+(1-\lambda)(y_1f_1(t)+\cdots+y_nf_n(t))$$

$$z_1f_1(t)+\cdots+z_nf_n(t)-f_0(t)=\lambda(x_1f_1(t)+\cdots+x_nf_n(t))-\lambda f_0(t)$$
$$+(1-\lambda)(y_1f_1(t)+\cdots+y_nf_n(t))-(1-\lambda)f_0(t)$$

So,

$$|z_1f_1(t)+\cdots+z_nf_n(t)-f_0(t)|\leq\lambda\,|x_1f_1(t)+\cdots+x_nf_n(t)-f_0(t)|$$
$$+(1-\lambda)\,|y_1f_1(t)+\cdots+y_nf_n(t)-f_0(t)|$$

Thus, this gives

$$|z_1f_1(t)+\cdots+z_nf_n(t)-f_0(t)|\leq\lambda\epsilon+(1-\lambda)\epsilon$$
$$|z_1f_1(t)+\cdots+z_nf_n(t)-f_0(t)|\leq\epsilon$$

So the linear combination $z_1f_1(t)+\cdots+z_nf_n(t)$ approximates $f_0(t)$ with tolerance ϵ, where $t\in[0,T]$ and $T=\min(T_1,T_2)$.

$$T_1\geq\alpha,\quad T_2\geq\alpha\Rightarrow T\geq\alpha$$

So,

$$W(x)\geq\alpha\Rightarrow x\in S_\alpha$$

$\therefore$ Since the sublevel set S_α is convex, $W(x)$ is quasiconcave.

Question 3.47

Suppose $f:\mathbb{R}^n\to\mathbb{R}$ is differentiable, $\operatorname{dom} f$ is convex, and $f(x)>0$ for all $x\in\operatorname{dom} f$. Show that f is log-concave if and only if for all $x,y\in\operatorname{dom} f$,

$$\frac{f(y)}{f(x)}\leq\exp\left(\frac{\nabla f(x)^T(y-x)}{f(x)}\right).$$

Solution:

Given, $f:\mathbb{R}^n\to\mathbb{R}$ is differentiable, $\operatorname{dom} f$ is convex, and $f(x)>0\quad\forall x\in\operatorname{dom} f$.
We need to show that,

$$f\text{ is log-concave}\quad\iff\quad\frac{f(y)}{f(x)}\leq\exp\left(\frac{\nabla f(x)^T(y-x)}{f(x)}\right)\quad\forall x,y\in\operatorname{dom} f$$

For a function $f(x)$ to be log-concave, $\log(f(x))$ should be concave on $\operatorname{dom} f$.
For concavity, $x,y\in\operatorname{dom} f$,

$$\log(f(y))\leq\log(f(x))+\nabla\log(f(x))^T(y-x)$$

The gradient of $\log(f(x))$ is

$$\nabla\log(f(x))=\frac{\nabla f(x)}{f(x)}$$

So, the condition is

$$\log(f(y)) \leq \log(f(x)) + \frac{\nabla f(x)^T(y-x)}{f(x)}$$

(i) If f is log-concave, then the inequality holds.
f is log-concave, so it will satisfy the inequality below:
For any two points $x, y \in \text{dom } f$,

$$\log(f(y)) \leq \log(f(x)) + \frac{\nabla f(x)^T(y-x)}{f(x)}$$

Thus,

$$\log\left(\frac{f(y)}{f(x)}\right) \leq \frac{\nabla f(x)^T(y-x)}{f(x)}$$

Applying exponential on both sides,

$$\frac{f(y)}{f(x)} \leq \exp\left(\frac{\nabla f(x)^T(y-x)}{f(x)}\right)$$

So, we have shown that if f is log-concave, the inequality

$$\frac{f(y)}{f(x)} \leq \exp\left(\frac{\nabla f(x)^T(y-x)}{f(x)}\right)$$

holds for all $x, y \in \text{dom } f$.

(ii) If the inequality holds, then f is log-concave.
Starting with the inequality,

$$\frac{f(y)}{f(x)} \leq \exp\left(\frac{\nabla f(x)^T(y-x)}{f(x)}\right)$$

Taking logarithm on both sides,

$$\log(f(y)) - \log(f(x)) \leq \frac{\nabla f(x)^T(y-x)}{f(x)}$$

This gives,

$$\log(f(y)) \leq \log(f(x)) + \frac{\nabla f(x)^T(y-x)}{f(x)}$$

This is the exact definition of concavity for the function $\log(f(x))$.
Since $\log(f(x))$ is concave, $f(x)$ is log-concave by definition.
Thus, we have shown both the directions:

$$\therefore \quad f \text{ is log-concave} \quad \iff \quad \frac{f(y)}{f(x)} \leq \exp\left(\frac{\nabla f(x)^T(y-x)}{f(x)}\right) \quad \forall x, y \in \text{dom } f$$

Question 3.49 (a)

Show that the following functions are log-concave.
(a) Logistic function: $f(x) = \frac{e^x}{1+e^x}$ with $\operatorname{dom} f = \mathbb{R}$.

Solution:

$$f(x) = \frac{e^x}{1+e^x}, \quad \operatorname{dom} f = \mathbb{R}$$

We need to show that $f(x)$ is log-concave.

$$\log(f(x)) = \log\left(\frac{e^x}{1+e^x}\right)$$

$$\log(f(x)) = x - \log(1+e^x)$$

For $\log(f(x))$ to be concave, we need its second derivative to be non-positive $\forall x \in \mathbb{R}$:

$$\frac{d}{dx}(\log(f(x))) = 1 - \frac{e^x}{1+e^x}$$

$$\frac{d}{dx}(\log(f(x))) = \frac{1}{1+e^x}$$

$$\frac{d^2}{dx^2}(\log(f(x))) = \frac{d}{dx}\left(\frac{1}{1+e^x}\right)$$

$$\frac{d^2}{dx^2}(\log(f(x))) = \frac{-e^x}{(1+e^x)^2}$$

Since $e^x > 0$ and $(1+e^x)^2 > 0$ for all $x \in \mathbb{R}$,
the second derivative of $\log(f(x))$ is always negative $\forall x \in \mathbb{R}$. So $\log(f(x))$ is concave.

$$\therefore f(x) = \frac{e^x}{1+e^x} \text{ is log-concave with domain } f = \mathbb{R}$$