Tutorial 4 and Assignment 2

March 8, 2024

Question 1

For each of the following functions determine whether it is convex, concave, quasiconvex, or quasiconcave.

(a) $f(x) = e^x - 1$ on $\mathbf{R}$.

(b) $f(x_1, x_2) = x_1 x_2$ on $\mathbf{R}^2_{++}$.

(c) $f(x_1, x_2) = 1/(x_1 x_2)$ on $\mathbf{R}^2_{++}$.

(d) $f(x_1, x_2) = x_1/x_2$ on $\mathbf{R}^2_{++}$.

(e) $f(x_1, x_2) = x_1^2/x_2$ on $\mathbf{R} \times \mathbf{R}_{++}$.

(f) $f(x_1, x_2) = x_1^\alpha x_2^{1-\alpha}$, where $0 \leq \alpha \leq 1$, on $\mathbf{R}^2_{++}$.

Question 2

Prove that $f(X) = \mathbf{tr}(X^{-1})$ is convex on $\mathbf{dom} f = \mathbf{S}^n_{++}$.

Solution:

Define $g(t) = f(Z + tV)$, where $Z > 0$ and $V \in S^n$.

$$
\begin{aligned}
g(t) = tr((Z + tV)^{-1}) &= tr\{Z^{-1}[I + tZ^{-1/2}VZ^{-1/2}]^{-1}\} \\
&= tr\{Z^{-1}Q(I + t\Lambda)^{-1}Q^T\} = tr\{Q^T Z^{-1} Q(I + t\Lambda)^{-1}\} \\
&= \sum_{i=1}^{n} (Q^T Z^{-1} Q)_{ii}(1 + t\lambda_i)^{-1},
\end{aligned}
$$

where we used the eigenvalue decomposition $Z^{-1/2}VZ^{-1/2} = Q\Lambda Q^T$. Noting that $(1 + t\lambda_i)^{-1}$ is convex, and $(Q^T Z^{-1} Q)_{ii} > 0$, hence $g(t)$ is convex.

Question 3

Show that for $p > 1$,

$$
f(x, t) = \frac{|x_1|^p + \cdots + |x_n|^p}{t^{p-1}} = \frac{\|x\|_p^p}{t^{p-1}}
$$

is convex on $\{(x, t) | t > 0\}$.

Solution:

This is the perspective function of $f(x) = ||x||_p^p = |x_1|^p + \ + |x_n|^p$.

(perspective function $g(x, t) = tf(x/t)$ is convex if f is convex).

Question 4

Derive the conjugates of the following functions.

(a) *Max function.* $f(x) = \max_{i=1,\dots,n} x_i$ on $\mathbf{R}^n$.

Solution: $f^*(y) = 0$ if $y > 0$ and $\sum y_i = 1$, otherwise $f^*(y) = +\infty$.

(try $n = 2, 3$ first, then prove it).

(b) *Piecewise-linear function on* $\mathbf{R}$. $f(x) = \max_{i=1,\dots,m} (a_i x + b_i)$ on $\mathbf{R}$. You can assume that the a_i are sorted in increasing order, *i.e.*, $a_1 \le \cdots \le a_m$, and that none of the functions $a_i x + b_i$ is redundant, *i.e.*, for each k there is at least one x with $f(x) = a_k x + b_k$.

Solution:

$$f^*(y) = -b_i - (b_{i+1} - b_i)\frac{y - a_i}{a_{i+1} - a_i}$$

if $y \in [a_i, a_{i+1}]$.

$$f^*(y) = +\infty$$

if $y < a_1$ or $y > a_m$.

(try $n = 2, 3$ first, then prove it).

(c) *Power function.* $f(x) = x^p$ on $\mathbf{R}_{++}$, where $p > 1$. Repeat for $p < 0$.

when $p > 1$, $f^*(y) = (p-1)(y/p)^q$ if $y > 0$, $= 0$ if $y \le 0$.

when $p < 0$, $f^*(y) = (p/q)(y/p)^q$ if $y \le 0$, $= +\infty$ if $y > 0$.

Question 5

Show that the conjugate of $f(X) = \mathbf{tr}\,(X^{-1})$ with $\mathbf{dom} f = \mathbf{S}^n_{++}$ is given by

$$f^*(Y) = -2\mathbf{tr}(-Y)^{1/2}, \quad \mathbf{dom} f^* = -\mathbf{S}^n_+.$$

Hint. The gradient of f is $\nabla f(X) = -X^{-2}$.

Pf:

- Suppose $Y \notin -S^n_+$ (i.e. the largest eigenvalue > 0. Do eigenvalue decomposition

$$Y = Q\Lambda Q^T = \sum \lambda_i q_i q_i^T$$

where $\lambda_1 > 0$. Let

$$X = Qdiag(t, 1, \dots, 1)Q^T = tq_1q_1^T + \sum_{i=2}^{n} q_i q_i^T$$

Then

$$tr(XY) - tr(X^{-1}) = [t\lambda_1 + \sum_{i=2}^{n} \lambda_i] - [1/t + (n-1)]$$

it $\to +\infty$ as $t \to +\infty$. So when $y \notin -S^n_+$, $f^*(Y) = +\infty$.

- When $Y \in -S^n_{++}$. Noting that

$$\nabla_X tr(XY) = Y, \nabla f(X) = -X^{-2}$$

To find the maximum of

$$trXY - trX^{-1}$$

by setting the gradient to zero, we obtain $X = (-Y)^{-1/2}$, and then

$$f^*(Y) = -2tr[(-Y)^{1/2}]$$

- When $Y \in -S^n_+$. Using the closeness of epigraph to handle it (not required).

Question 6

Conjugate of negative normalized entropy. Show that the conjugate of the negative normalized entropy

$$f(x) = \sum_{i=1}^{n} x_i \log\left(x_i/\mathbf{1}^T x\right),$$

with $\mathbf{dom} f = \mathbf{R}^n_{++}$ is given by

$$f^*(y) = \begin{cases} 0 & \sum_{i=1}^n e^{y_i} \leq 1 \\ +\infty & \text{otherwise.} \end{cases}$$

Solution:

- When $\sum_{i=1}^n e^{y_i} > 1$, Let $x_i = te^{y_i}$.

$$y^T x - f(x) = t\sum_{i=1}^{n} e^{y_i} log(\sum_{k=1}^{n} e^{y_k})$$

  $\rightarrow +\infty$ as $t \rightarrow +\infty$. Thus $f^*(y) = +\infty$.

- When $\sum_{i=1}^n e^{y_i} \leq 1$, Let $z_i = e^{y_i}$. Then

$$y^T x - f(x) = \sum_{i=1}^{n} x_i log[\frac{z_i(\sum_{k=1}^n x_k)}{x_i}] = -D_{kl}(x, (\sum x_k)z) + (\sum_k x_k)(1 - \sum z_j)$$

  Noting that $\sum z_j \leq 1$ and $D_{kl} \geq 0$, $y^T x - f(x) \leq 0$ for all $x > 0$. Moreover, $y^T x - f(x) \rightarrow 0$ if $x = (t, \ldots, t)$ and $t \rightarrow 0$. So:

$$f^*(x) = sup_{x>0}[y^T x - f(x)] = 0$$