Assignment #3

Shao Hua, Huang 0750727
ECM5901 - Optimization Theory and Application

December 27, 2019

Exercise 1. (Kullback-Leibler divergence and the information inequality.) The Kullback-Leibler divergence between $u, v \in \mathbb{R}^n_{++}$ is defined as

$$D_{kl}(u,v) = \sum_{i=1}^{n} (u_i \log(u_i/v_i) - u_i + v_i)$$

Show that

(a) $D_{kl}(u,v)$ is convex (5%)

(b) $D_{kl}(u,v) \geq 0$ for all $u, v \in \mathbb{R}^n_{++}$ (5%)

(c) $D_{kl}(u,v) = 0$ if and only if $u = v$ (5%)

Hint: $D_{kl}(u,v)$ can be expressed as

$$D_{kl}(u,v) = f(u) - f(v) - \nabla f(v)^T (u - v) \tag{1}$$

where $f(v) = \sum_{i=1}^{n} v_i \log v_i$ is the nagative entropy of v for $v \in \mathbb{R}^n_{+}$.

Proof.

(a) The *negative logarithm* function $f(x) = -\log x$ on $\mathbb{R}_{++}$ is (strictly) convex.
Apply §3.2.6, perspective of $f(x)$, we get *relative entropy* function

$$g_1(x,t) = t \log(t/x) \tag{2}$$

is (strictly) convex.
Let $g_2(x,t) = -t + x$, then g_2 is an affine function (both convex and concave).
Apply §3.2.1, nonnegative weighted sums of $g_1(x,t)$ and $g_2(x,t)$, we get

$$h(x,t) = g_1(x,t) + g_2(x,t) = t \log(t/x) - t + x$$

is also convex.
Apply §3.2.1 again, we can extends $h(x,t)$ to dimension n, therefore

$$D_{kl}(u,v) = \sum_{i=1}^{n} (u_i \log(u_i/v_i) - u_i + v_i)$$

is convex.

(b) Substitute x with 1 in (2) and extends the function to dimension n by §3.2.1, we know that the *negative entropy* function

$$f(v) = \sum_{i=1}^{n} v_i \log v_i$$

is strictly convex. Therefore, for $u, v \in \mathbf{dom}\, f$ and $u \neq v$ we have

$$\begin{aligned}
&f(u) > f(v) + \nabla f(v)^T (u - v) \\
\Rightarrow &\sum_{i=1}^{n} u_i \log u_i > \sum_{i=1}^{n} v_i \log v_i + \sum_{i=1}^{n} (\log v_i + 1)(u_i - v_i) \\
\Rightarrow &\sum_{i=1}^{n} u_i \log u_i > \sum_{i=1}^{n} u_i \log v_i + \sum_{i=1}^{n} (u_i - v_i) \\
\Rightarrow &\sum_{i=1}^{n} (u_i \log(u_i/v_i) - u_i + v_i) > 0 \\
\Rightarrow &D_{kl}(u, v) > 0
\end{aligned}$$

(c) $u = v \Rightarrow D_{kl}(u, v) = D_{kl}(u, u) = \sum_{i=1}^{n} (u_i \log(u_i/u_i) - u_i + u_i) = 0$

$D_{kl}(u, v) = 0$ with $u \neq v$ contradicts to result of (b), *i.e.*, $D_{kl}(u, v) = 0 \Rightarrow u = v$
Therefore, we have $u = v \Leftrightarrow D_{kl}(u, v) = 0$ $\square$

Exercise 2. Adapt the proof of concavity of the log-determinant function in §3.1.5 of [BV04] to show that

$$f(X) = \mathbf{tr}(X^{-1})$$

is convex on $\mathbf{dom}\, f = \mathbb{S}_{++}^{n}$. (15%)

Proof. Consider an arbitrary line, given by $X = Z + tV$ where $Z, V \in \mathbb{S}^n$. Define $g(t) = f(Z + tV)$, and restrict g to the interval of values of t for which $Z + tV \succ 0$. Without loss of generality, we can assume that $t = 0$ is inside this interval, *i.e.*, $Z \succ 0$. We have

$$\begin{aligned}
g(t) &= \mathbf{tr}((Z + tV)^{-1}) \\
&= \mathbf{tr}(Z^{-1}(I + tZ^{-\frac{1}{2}} V Z^{-\frac{1}{2}})^{-1}) \\
&= \mathbf{tr}(Z^{-1}(I + tQ\Lambda Q^T)^{-1}) \\
&= \mathbf{tr}(Z^{-1} Q (I + t\Lambda)^{-1} Q^T) \\
&= \mathbf{tr}(Q^T Z^{-1} Q (I + t\Lambda)^{-1}) \\
&= \sum_{i=1}^{n} (Q^T Z^{-1} Q)_{ii} (1 + t\lambda_i)^{-1}
\end{aligned}$$

where $Z^{-\frac{1}{2}} V Z^{-\frac{1}{2}} = Q\Lambda Q^T$ is the eigenvalue decomposition.
Therefore, g is nonnegative weighted sums of convex functions $(1+t\lambda_i)^{-1}$, *i.e.*, g is convex. $\square$

Exercise 3. (Composition rules.) Show that the following functions are convex

(a) $f(x,u,v) = -\log(uv - x^T x)$ on
$\mathbf{dom}\, f = \{(x,u,v) \mid x \in \mathbb{R}^n,\ u,v \in \mathbb{R},\ uv > x^T x,\ u,v > 0\}$

(b) Show that

$$f(x) = \frac{\|Ax+b\|_2^2}{c^T x + d}$$

is convex on $\{x \in \mathbb{R}^n \mid c^T x + d > 0\}$, where $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m, c \in \mathbb{R}^n$ and $d \in \mathbb{R}$. (10%)

Hint: $x^T x/u$ is convex in (x,u) for $u > 0$.

Proof.

(a) Rewrite $f(x,u,v) = -\log u - \log(v - x^T x/u)$. We know that $-\log u$ is convex, and $v - x^T x/u$ is concave because v linear and $x^T x/u$ convex on $\{(x,u) \mid u > 0\}$.
By applying the rule of composition on $f(x) = h(g(x))$:

$$f \text{ is convex if } h \text{ is convex and nonincreasing, and } g \text{ is concave}$$

We get $-\log(v - x^T x/u)$ is also convex.
Therefore, $f(x,u,v)$ is convex (nonnegative weighted sums of convex functions).

(b) The function is composed by the convex function $g(y,t) = y^T y/t$ and the affine mapping $f(x) = g(Ax+b, c^T x + d)$. From §3.2.2 and $g(y,t)$ with $t > 0$ is convex, we know that $f(x)$ is also convex. □

Exercise 4. (Conjugate of convex plus affine function) Define $g(x) = f(x) + c^T x + d$ for $c \in \mathbb{R}^n, d \in \mathbb{R}$, where $f : \mathbb{R}^n \to \mathbb{R}$ is convex. Express g^* in terms of f^* and c, d. (10%)

Solution.

$$\begin{aligned} g^*(y) &= \sup_{x \in \mathbf{dom}\, f} (y^T x - f(x) - c^T x - d) \\ &= \sup_{x \in \mathbf{dom}\, f} ((y-c)^T x - f(x)) - d \\ &= f^*(y-c) - d \end{aligned}$$

Exercise 5. (Log-concavity of Gaussian cumulative distribution function.) The cumulative distribution function of a Gaussian random variable,

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt,$$

is log-concave. (This follows from the general result that the convolution of two log-concave functions is log-concave.) In this problem we guide you through a simple self-contained proof that f is log-concave. Recall that f is log-concave if and only if $f''(x)f(x) \le f'(x)^2$ for all x.

(a) Verify that $f''(x)f(x) \leq f'(x)^2$ for $x \geq 0$. That leaves us the hard part, which is to show the inequality for $x < 0$.

(b) Verify that for any t and x, we have $t^2/2 \geq -x^2/2 + xt$. (Hint: consider the first order condition with the function $g(t) = t^2/2$.)

(c) Use part (b) to show that $e^{-t^2/2} \leq e^{x^2/2 - xt}$. Conclude that for $x < 0$,

$$\int_{-\infty}^{x} e^{-t^2/2} dt \leq e^{x^2/2} \int_{-\infty}^{x} e^{-xt} dt.$$

(d) Use part (c) to verify that $f''(x)f(x) \leq f'(x)^2$ for $x \leq 0$.

(20%)

Proof.

(a) $f(x) > 0$, $f'(x) = \dfrac{e^{-x^2/2}}{\sqrt{2\pi}} > 0$, and $f''(x) = -\dfrac{xe^{-x^2/2}}{\sqrt{2\pi}} \Rightarrow f''(x) \leq 0$ for $x \geq 0$.
Hence, $f''(x)f(x) \leq f'(x)^2$ for $x \geq 0$

(b) $t^2/2$ is convex on $\mathbb{R}$, employ first-order condition, we have

$$\frac{t^2}{2} \geq \frac{x^2}{2} + x(t - x) = -\frac{x^2}{2} + xt \tag{3}$$

The equation holds for any $x, t \in \mathbb{R}$.

(c) Take exponentials of (3), we have $e^{-t^2/2} \leq e^{x^2/2 - xt}$.
Then take integrate, we have

$$\begin{aligned} \int_{-\infty}^{x} e^{-t^2/2} dt &\leq e^{x^2/2} \int_{-\infty}^{x} e^{-xt} dt \\ &= e^{x^2/2} \left(-\frac{e^{-x^2}}{x} \right) = -\frac{e^{-x^2/2}}{x} \end{aligned} \tag{4}$$

(d) The inequality $f''(x)f(x) \leq f'(x)^2$ becomes

$$\begin{aligned} &\left(-\frac{xe^{-x^2/2}}{\sqrt{2\pi}} \right) \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt \right) \leq \left(\frac{e^{-x^2/2}}{\sqrt{2\pi}} \right)^2 \\ \Rightarrow &- xe^{-x^2/2} \int_{-\infty}^{x} e^{-t^2/2} dt \leq e^{-x^2} \\ \Rightarrow &\int_{-\infty}^{x} e^{-t^2/2} dt \leq -\frac{e^{-x^2/2}}{x} \quad \text{(when } x < 0\text{)} \end{aligned} \tag{5}$$

(4) and (5) are actually the same inequality.
Therefore, $f''(x)f(x) \leq f'(x)^2$ holds for $x \leq 0$ (it also holds trivially when $x = 0$). $\square$

Exercise 6. (Sublevel sets and epigraph of K-convex functions.) Let $K \subset \mathbb{R}^m$ be a proper convex cone with associated generalized inequality K, and let $f : \mathbb{R}^n \to \mathbb{R}^m$. For $\alpha \in \mathbb{R}^m$, the α-sublevel set of f (with respect to K) is defined as

$$C_\alpha = \{x \in \mathbb{R}^n \mid f(x) \preceq_K \alpha\}.$$

The epigraph of f, with respect to K, is defined as the set

$$\underset{K}{\mathbf{epi}}\, f = \{(x,t) \in \mathbb{R}^{n+m} \mid f(x) \preceq_K t\}$$

Show the following:

(a) If f is K-convex, then its sublevel sets C_α are convex for all α. (10%)

(b) f is K-convex if and only if $\underset{K}{\mathbf{epi}}\, f$ is a convex set. (10%)

Proof.

(a) For all $x, y \in C_\alpha$, and $0 \leq \theta \leq 1$,

$$\begin{aligned} f(\theta x + (1-\theta y)) &\preceq_K \theta f(x) + (1-\theta) f(y) && (f \text{ is } K\text{-convex}) \\ &\preceq_K \theta\alpha + (1-\theta)\alpha = \alpha && (x, y \in C_\alpha) \end{aligned}$$

Therefore, C_α is convex.

(b) Proposition of $\to$, $(x,u), (y,v) \in \underset{K}{\mathbf{epi}}\, f$, and $0 \leq \theta \leq 1$,

$$\begin{aligned} f(\theta x + (1-\theta y)) &\preceq_K \theta f(x) + (1-\theta) f(y) && (f \text{ is } K\text{-convex}) \\ &\preceq_K \theta u + (1-\theta) v && ((x,u),(y,v) \in \underset{K}{\mathbf{epi}}\, f) \end{aligned}$$

$(\theta x + (1-\theta y), \theta u + (1-\theta v)) \in \underset{K}{\mathbf{epi}}\, f$, *i.e.*, $\underset{K}{\mathbf{epi}}\, f$ is convex.
Proposition of $\leftarrow$, let $f(x) = u, f(y) = v$ $((x,u),(y,v) \in \underset{K}{\mathbf{epi}}\, f)$, and $0 \leq \theta \leq 1$,

$$\begin{aligned} f(\theta x + (1-\theta y)) &\preceq_K \theta u + (1-\theta) v && (\underset{K}{\mathbf{epi}}\, f \text{ is convex}) \\ &= \theta f(x) + (1-\theta) f(y) \end{aligned}$$

We get $\underset{K}{\mathbf{epi}}\, f$ is K-convex.
Therfore, f is K-convex if and only if $\underset{K}{\mathbf{epi}}\, f$ is convex. □