HW3 Solutions

October 31, 2019

1. (a) Suppose $(u, v), (t, s) \in \operatorname{dom} D_{kl}$.

$$\begin{aligned}
&D_{kl}(\theta u + (1-\theta)t, \theta v + (1-\theta)s) \\
&= \sum_{i=1}^{n} [(\theta u_i + (1-\theta)t_i) \log \frac{(\theta u_i + (1-\theta)t_i)}{(\theta v_i + (1-\theta)s_i)} - (\theta u_i + (1-\theta)t_i) + (\theta v_i + (1-\theta)s_i)] \\
&\leq \sum_{i=1}^{n} [\theta u_i \log \frac{\theta u_i}{\theta v_i} + (1-\theta)t_i \log \frac{\theta t_i}{\theta s_i} - (\theta u_i + (1-\theta)t_i) + (\theta v_i + (1-\theta)s_i)] \\
&\text{(by log sum inequality)} \\
&= \theta D_{kl}(u, v) + (1-\theta) D_{kl}(t, s).
\end{aligned}$$

$\Rightarrow D_{kl}(u, v)$ is convex.

(b) The negative entropy is strictly convex and differentiable on $\mathbb{R}^n_{++}$, hence

$$f(u) > f(v) + \nabla f(v)^T (u - v)$$

for all $u, v \in \mathbb{R}^n_{++}$ with $u \neq v$. Evaluating both sides of the inequality, we obtain

$$\begin{aligned}
\sum_{i=1}^{n} u_i \log u_i &> \sum_{i=1}^{n} v_i \log v_i + \sum_{i=1}^{n} (\log v_i + 1)(u_i - v_i) \\
&= \sum_{i=1}^{n} u_i \log v_i + \mathbf{1}^T (u - v).
\end{aligned}$$

Rearranging this inequality, we have $D_{kl}(u, v) > 0$.

Otherwise, if u=v, $D_{kl}(u, v) = 0$.

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

($\Leftarrow$) If $D_{kl}(u, v) = 0$, that indicate $u_i \log(u_i/v_i) = u_i - v_i$

$$\Rightarrow \log(u_i/v_i) = 1 - (v_i/u_i), \forall i.$$

Let $x_i = (u_i/v_i)$.

$$\log(x_i) = 1 - (1/x_i).$$

Equality holds if $x_i = 1$ (by fundamental inequality)

$\Rightarrow u = v$.

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

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

where we used the eigenvalue decomposition $Z^{-1/2}VZ^{-1/2} = Q\Lambda Q^T$. In the last equality we express g as a positive weighted sum of convex functions $1/(1 + t\lambda_i)$. Hence it is convex.

3. (a) We can express f as

$$f(x, u, v) = -\log u - \log(v - x^T x/u).$$

The first term is convex. The function $v - x^T x/u$ is concave because v is linear and $x^T x/u$ is convex on $\{(x, u) \mid u > 0\}$. Therefore the second term in f is convex: it is the composition of a convex decreasing function $-\log t$ and a concave function.

(b) Let $g(y, t) = y^T y/t$ is convex for $t > 0$.

$$f(x) = g(Ax + b, c^T x + d)$$

is the composition with the affine mapping.

$\Rightarrow f(x)$ is convex.

4.

$$\begin{aligned} g^*(y) &= \sup(y^T x - f(x) - c^T x - d) \\ &= \sup((y - c)^T x - f(x)) - d \\ &= f^*(y - c) - d. \end{aligned}$$

5. The derivatives of f are

$$f'(x) = e^{-x^2/2}/\sqrt{2\pi}, \qquad f''(x) = -xe^{-x^2/2}/\sqrt{2\pi}.$$

(a) $f''(x) \le 0$ for $x \ge 0$.

(b) Since $t^2/2$ is convex we have

$$t^2/2 \ge x^2/2 + x(t - x) = xt - x^2/2.$$

This is the general inequality

$$g(t) \ge g(x) + g'(x)(t - x),$$

which holds for any differentiable convex function, applied to $g(t) = t^2/2$.

(c) Take exponentials and integrate from (b).

$$e^{-t^2/2} \leq e^{x^2/2 - xt}$$

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

(d) This basic inequality reduces to

$$-xe^{-x^2/2} \int_{-\infty}^{x} e^{-t^2/2} dt \leq e^{-x^2},$$

i.e.,

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

This follows from part (c) because

$$\int_{-\infty}^{x} e^{-xt} dt = \frac{e^{-x^2}}{-x}.$$

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

$$f(\theta x + (1-\theta)y) \preceq_K \theta f(x) + (1-\theta) f(y) \preceq_K \alpha.$$

(b)($\Rightarrow$) If f is K-convex, for any $(x, u), (y, v) \in \textbf{epi}\, f$, and $0 \leq \theta \leq 1$,

$$f(\theta x + (1-\theta)y) \preceq_K \theta f(x) + (1-\theta) f(y) \preceq_K \theta u + (1-\theta) v$$

$\Rightarrow$ $\textbf{epi}\, f$ is convex set.

($\Leftarrow$) If $\textbf{epi}\, f$ is convex set, for any $(x, u), (y, v) \in \textbf{epi}\, f \Rightarrow \theta(x, u) + (1-\theta)(y, v) \in \textbf{epi}\, f$

Then,

$$f(\theta x + (1-\theta)y) \preceq_K \theta u + (1-\theta) v.$$

If we choose $f(x) = u, f(y) = v$

$$f(\theta x + (1-\theta)y) \preceq_K \theta f(x) + (1-\theta) f(y).$$

$\Rightarrow$ f is K-convex.