Yale University
CPSC 516, Spring 2023
Assignment 3

Chang Feng (Felix) Zhou cz397

P.1.

(a)

The gradient and Hessian are given by

$$\begin{aligned}
\nabla f(x) &= \log x \\
\nabla^2 f(x) &= \mathrm{Diag}(1/x_1, \dots, 1/x_n).
\end{aligned}$$

Here the logarithm is applied component-wise and the Diag : $\mathbb{R}^n \to \mathbb{R}^{n\times n}$ maps a vector to the diagonal matrix whose non-zero entries are precisely given by the input vector.

(b)

By our work in class, f is strictly convex if and only if $\nabla^2 f \succ 0$. Indeed, the eigenvalues of $\nabla^2 f$ are given by

$$1/x_1, \dots, 1/x_n$$

and are all positive over $\mathbb{R}^n_{>0}$. Thus f is indeed strictly convex.

(c)

Suppose towards a contradiction that f is α-strongly convex for some $\alpha > 0$. Let e_1 be the vector which is all zero except for a one in the first entry. We have for all $\lambda > 0$,

$$\begin{aligned}
f(\lambda e_1 + \lambda e_1) &\geq f(\lambda e_1) + \langle \nabla f(\lambda e_1), \lambda e_1 \rangle + \frac{\alpha}{2}\|\lambda e_1\|_2^2 \\
2\lambda \log(2\lambda) - 2\lambda &\geq \lambda \log \lambda - \lambda + \lambda \log \lambda + \frac{\alpha}{2}\lambda^2 \\
2\lambda \log\left(\frac{2\lambda}{\lambda}\right) &\geq \lambda + \frac{\alpha}{2}\lambda^2 \\
2\log 2 &\geq 1 + \frac{\alpha}{2}\lambda.
\end{aligned}$$

This is a contradiction since we can make the RHS arbitrarily large while the LHS stays constant.

(d)

$$\begin{aligned}
D_f(x,y) &:= f(y) - f(x) - \langle \boldsymbol{\nabla} f(x), y - x\rangle \\
&= \sum_{i=1}^n y_i \log y_i - y_i - x_i \log x_i + x_i - (y_i - x_i)\log x_i \\
&= \sum_i y_i \log y_i - y_i + x_i - y_i \log x_i \\
&= \boxed{\sum_{i=1}^n y_i \log \frac{y_i}{x_i} + x_i - y_i}.
\end{aligned}$$

By inspection, this is not symmetric for all $x, y > 0$.

(e)

We wish to show that

$$D_f(x,y) \geq \frac{1}{2}\|y - x\|_1^2$$

for all $x, y \in \Delta^n := \{x \in \mathbb{R}^n_{>0} : \sum_{i=1}^n x_i = 1\}$. First, we remark that in Δ^n,

$$D_f(x,y) = \sum_{i=1}^n y_i \log \frac{y_i}{x_i} = \mathrm{KL}(y\|x).$$

Now, Pinsker's inequality states that

$$D_f(x,y) = \mathrm{KL}(y\|x) \geq \frac{1}{2}\|y - x\|_1^2.$$

and so f is 1-strongly convex with respect to the 1-norm, as desired.

P.2.

(a)

(b)

(c)

(d)

(e)