CSE 291-A: Homework 2

Hao-en Sung [A53204772] (wrangle1005@gmail.com)
Department of Computer Science
University of California, San Diego
San Diego, CA 92092

Exercises from textbook Chapter Three

3.2 Level sets of convex, concave, quasiconvex, and quasiconcave functions.

Sol. For the first curve, it is at least quasiconvex, because the curve label increases when the curve getting larger. However, it is not convex, since the space between curves does not always decrease with the increase of curve label.

For the second curve, it is at least quasiconcave, because the curve label decreases when the curve getting larger. At the same time, it is also concave, since the space between curves monotonically decrease with the increase of curve label. □

3.13 Kullback-Leibler divergence and the information inequality.

Sol. Firstly, I can prove that $f(v_i) = v_i \log v_i$ is strictly convex by calculating its second derivative, i.e. $f''(v_i) = \frac{1}{v_i} > 0$ when **dom** $v_i \in \mathbb{R}_{++}$. Thus, $f(v) = \sum_{i=1}^{n} v_i \log v_i$ is also strictly convex, since it is basically the summation of convex functions.

Later, based on the convex function property, I know that

$$\begin{aligned} f(u) &\geq f(v) + \Delta f(v)^\intercal (u - v) \\ f(u) - f(v) - \Delta f(v)^\intercal (u - v) &\geq 0 \\ D_{kl} &\geq 0 \end{aligned}$$

.

On top of that, since $f(v)$ is strictly convex, the equal sign happens only when $u = v$. □

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

(a) $f(x) = e^x - 1,\ x \in \mathbb{R}$

Sol. By calculating the second derivative of $f(x)$:

$$\begin{aligned} f'(x) &= e^x \\ f''(x) &= e^x \geq 0, \end{aligned}$$

I know f(x) is convex, which also implies quasiconvex. □

(b) $f(x_1, x_2) = x_1 x_2,\ (x_1, x_2) \in \mathbb{R}^2_{++}$

Sol. By calculating the second derivative of $f(x)$:

$$\begin{aligned} f'(x_1, x_2) &= \begin{bmatrix} x_2 & x_1 \end{bmatrix} \\ f''(x_1, x_2) &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \end{aligned}$$

I know that it is neither convex or concave, since $0 \cdot 0 - 1 \cdot 1 < 0$ implies it is neither positive semi-definite nor negative semi-definite.

It is obvious that $\{(x_1, x_2) | x_1 x_2 \geq \alpha\}$ is a convex set. Thus, it is quasiconcave but not quasiconvex. $\square$

(c) $f(x_1, x_2) = \frac{1}{x_1 x_2},\ (x_1, x_2) \in \mathbb{R}^2_{++}$

Sol. By calculating the second derivative of $f(x)$:

$$f'(x_1, x_2) = \begin{bmatrix} \frac{-1}{x_1^2 x_2} & \frac{-1}{x_1 x_2^2} \end{bmatrix}$$
$$f''(x_1, x_2) = \frac{1}{x_1 x_2} \begin{bmatrix} \frac{2}{x_1^2} & \frac{1}{x_1 x_2} \\ \frac{1}{x_1 x_2} & \frac{2}{x_2^2} \end{bmatrix},$$

I know $f(x_1, x_2)$ is both a convex and quasiconvex function, since

$$\frac{2}{x_1^2} \geq 0$$
$$\frac{2}{x_2^2} \geq 0$$
$$\frac{2}{x_1^2} \cdot \frac{2}{x_2^2} - \frac{1}{x_1 x_2} \cdot \frac{1}{x_1 x_2} = \frac{4-1}{x_1^2 x_2^2} \geq 0.$$

Slightly different from (b), $\{(x_1, x_2) | x_1 x_2 \leq \alpha\}$ is a convex set; thus, it is neither quasiconcave nor concave. $\square$

(d) $f(x_1, x_2) = \frac{x_1}{x_2},\ (x_1, x_2) \in \mathbb{R}^2_{++}$

Sol. By calculating the second derivative of $f(x)$:

$$f'(x_1, x_2) = \begin{bmatrix} \frac{1}{x_2} & \frac{-x_1}{x_2^2} \end{bmatrix}$$
$$f''(x_1, x_2) = \begin{bmatrix} 0 & \frac{-1}{x_2^2} \\ \frac{-1}{x_2^2} & \frac{2x_1}{x_2^3} \end{bmatrix},$$

I know that it is neither convex or concave, since $0 \cdot \frac{2x_1}{x_2^3} - \frac{-1}{x_2^2} \cdot \frac{-1}{x_2^2} = -\frac{1}{x_2^4} < 0$ implies it is neither positive semi-definite nor negative semi-definite.

It is obvious that either $\{(x_1, x_2) | \frac{x_1}{x_2} \leq \alpha\}$ or $\{(x_1, x_2) | \frac{x_1}{x_2} \geq \alpha\}$ represents a halfspace. Thus, it is quasilinear (both quasiconvex and quasiconcave). $\square$

(e) $f(x_1, x_2) = \frac{x_1^2}{x_2},\ (x_1, x_2) \in \mathbb{R} \times \mathbb{R}_{++}$

Sol. By calculating the second derivative of $f(x)$:

$$f'(x_1, x_2) = \begin{bmatrix} \frac{2x_1}{x_2} & \frac{-x_1^2}{x_2^2} \end{bmatrix}$$
$$f''(x_1, x_2) = \begin{bmatrix} \frac{2}{x_2} & \frac{-2x_1}{x_2^2} \\ \frac{-2x_1}{x_2^2} & \frac{2x_1^2}{x_2^3} \end{bmatrix},$$

I know $f(x_1, x_2)$ is both a convex and quasiconvex function, since

$$\frac{2}{x_2} \geq 0$$
$$\frac{2x_1^2}{x_2^3} \geq 0$$
$$\frac{2}{x_2} \cdot \frac{2x_1^2}{x_2^3} - \frac{-2x_1}{x_2^2} \cdot \frac{-2x_1}{x_2^2} = 0.$$

One can find out $\{(x_1, x_2)|\frac{x_1^2}{x_2} \geq \alpha\}$ represents the area below a parabola, which is not convex. Thus, this function is neither quasiconcave nor concave. $\square$

(f) $f(x_1, x_2) = x_1^\alpha x_2^{1-\alpha},\ 0 \leq \alpha \leq 1,\ (x_1, x_2) \in \mathbb{R}^2_{++}$

Sol. By calculating the second derivative of $f(x)$:

$$f'(x_1, x_2) = \begin{bmatrix}\alpha \cdot x_1^{\alpha-1} \cdot x_2^{1-\alpha} & (1-\alpha)x_1^\alpha x_2^{-\alpha}\end{bmatrix}$$
$$f''(x_1, x_2) = \begin{bmatrix}\alpha(\alpha-1)x_1^{\alpha-2}x_2^{1-\alpha} & \alpha(1-\alpha)x_1^{\alpha-1}x_2^{-\alpha} \\ \alpha(1-\alpha)x_1^{\alpha-1}x_2^{-\alpha} & (1-\alpha)(-\alpha)x_1^\alpha x_2^{-1-\alpha}\end{bmatrix},$$

I know $f(x_1, x_2)$ is both a concave and quasiconcave function, since

$$\alpha(\alpha-1)x_1^{\alpha-2}x_2^{1-\alpha} \leq 0$$
$$(1-\alpha)(-\alpha)x_1^\alpha x_2^{-1-\alpha} \leq 0$$
$$\alpha(\alpha-1)x_1^{\alpha-2}x_2^{1-\alpha} \cdot (1-\alpha)(-\alpha)x_1^\alpha x_2^{-1-\alpha} - \alpha(1-\alpha)x_1^{\alpha-1}x_2^{-\alpha} \cdot \alpha(1-\alpha)x_1^{\alpha-1}x_2^{-\alpha} = 0.$$

Similar to (b), it is neither quasiconvex nor convex. $\square$

3.24 Some functions on the probability simplex.

(a) $\mathbf{E}x$

Sol. It can be rewritten as $\mathbf{E}x = \sum_{i=1}^n p_i \cdot a_i$. Since the summation of linear function is still linear, it is convex, quasiconvex, concave, and quasiconcave. $\square$

(b) $\mathbf{prob}(x \geq \alpha)$

Sol. Since it will sum over only a subset of all linear functions, which is still linear, it is convex, quasiconvex, concave, and quasiconcave. $\square$

(c) $\mathbf{prob}(\alpha \leq x \leq \beta)$

Sol. Similar to (b), summation of a subset of all linear functions is still linear, and thus, it is convex, quasiconvex, concave, and quasiconcave. $\square$

(d) $\sum_{i=1}^n p_i \log p_i$

Sol. It is known that $p_i \log p_i$ is a convex function. Thus, the summation of n convex functions are still convex (and quasiconvex).

If one draws the figure for $p_i > 0$, one can tell that $\{p_i | p_i \log p_i \geq \alpha\}$ is convex. Thus, it is quasiconcave. However, it is not concave, since it is strictly convex. $\square$

(e) $\mathbf{var}x = \mathbf{E}(x - \mathbf{E}x)^2$

Sol. One can rewrite the variance as

$$\begin{aligned}\mathbf{var}x &= \mathbf{E}(x^2) - \mathbf{E}(x)^2 \\ &= \sum_{i=1}^n p_i \cdot a_i^2 - \left(\sum_{i=1}^n p_i \cdot a_i\right)^2,\end{aligned}$$

which is a quadratic function for p with leading negative sign. In other words, $\mathbf{var}x$ is both concave and quasiconcave.

To verify whether this function is not quasiconvex (and thus not convex), I need to draw a proper example with $n = 2$. For example, let $a_0 = 0, a_1 = 1$. No matter $p_0 = \frac{1}{4}, p_1 = \frac{3}{4}$ or, $p_0 = \frac{3}{4}, p_1 = \frac{1}{4}$, I have $\mathbf{var}x = \frac{3}{16}$. However, when $p_0 = \frac{1}{2}, p_1 = \frac{1}{2}$, I have $\mathbf{var}x = \frac{1}{4} < \frac{3}{16}$. Since its suplevel set is not convex, it is neither quasiconvex nor convex. $\square$

(f) $\textbf{quartile}(x) = \textbf{inf}\{\beta \mid \textbf{prob}(x \le \beta) \ge 0.25\}$

Sol. Since $\textbf{quartile}(x)$ can only output $\{a_1, \cdots, a_n\}$, which is not continuous, it is neither convex nor concave.

On the other hand, both $\{x|\textbf{quartile}(x) \le 0\}$ and $\{x|\textbf{quartile}(x) \ge 0\}$ are convex sets. Thus, it is quasilinear (both quasiconvex and quasiconcave). $\square$

Assignments

1. [log-sum-exp]

Sol. Firstly, I am going to verify that $g(x) = \log\left(\sum_i e^{x_i}\right)$ is a convex function.

$$\frac{\partial g(x)}{\partial x_i} = \frac{e^{x_i}}{\sum_i e^{x_i}}$$

The calculation of the second derivative of $g(x)$ needs to consider two different cases.

$$\frac{\partial g(x)}{\partial^2 x_i} = \frac{e^{x_i} \cdot \sum_i e^{x_i} - e^{x_i} \cdot e^{x_i}}{\left(\sum_i e^{x_i}\right)^2}$$
$$\frac{\partial g(x)}{\partial x_i x_j} = \frac{-e^{x_i} \cdot e^{x_j}}{\left(\sum_i e^{x_i}\right)^2}$$

Assume $z_i = e^{x_i}$, the second derivative of $g(x)$ can be rewritten as follows.

$$\frac{\partial g(x)}{\partial^2 x} = \frac{\text{diag}(z)\mathbf{1}^\intercal z - zz^\intercal}{\left(\sum_i z_i\right)^2}$$

For arbitrary vector v, we have

$$v^\intercal \frac{\partial g(x)}{\partial^2 x} v = \frac{\sum_i z_i v_i^2 \cdot \sum_i z_i - \left(\sum_i v_i z_i\right)^2}{\left(\sum_i z_i\right)^2} \ge 0,$$

according to Cauchy-Schwarz inequality. Thus, $g(x)$ is convex.

Similarly, I can also verify that $h(x) = \log\left(\sum_i e^{-x_i}\right)$ is a convex function.

$$\frac{\partial h(x)}{\partial x_i} = \frac{-e^{-x_i}}{\sum_i e^{-x_i}}$$

Assume $z_i' = e^{-x_i}$, the second derivative of $h(x)$ can be rewritten as follows.

$$\frac{\partial h(x)}{\partial^2 x} = \frac{\text{diag}(z')\mathbf{1}^\intercal z' - z' {z'}'^\intercal}{\left(\sum_i z_i'\right)^2}$$

Since it has almost the same form as the second derivative of $g(x)$, I can also prove that $\frac{\partial h(x)}{\partial^2 x}$ is positive semi-definite, and thus, $h(x)$ is convex.

It is known that the summation of convex functions are still convex. In other words, $f(x) = g(x) + h(x)$ is convex. $\square$

2. [Least Action]

(a)

The proof that the least action principle implying a concave function can be found in The Principle of Least Action (Caltech).

(b)

One can derive the force from the energy function:

$$
\begin{aligned}
U(x) &= \frac{1}{m}\int_{\infty}^{x} F(x)\cdot dx \\
&= \frac{1}{m}\int_{\infty}^{x} \frac{GMm}{x^2}\cdot dx \\
&= -\frac{GM}{x} \\
\frac{\partial U(x)}{\partial x} &= \frac{1}{m}F(x) = U'(x) \\
F(x) &= m\cdot U'(x) \\
&= V'(x) = -\frac{k}{x^2}
\end{aligned}
$$

With the derivation, I know that the force $F(x)$ for $1 \leq x \leq 10$ are as follows.

$$
\begin{aligned}
F(1) &\approx -0.010000000000000 \\
F(2) &\approx -0.002500000000000 \\
F(3) &\approx -0.001111111111111 \\
F(4) &\approx -0.000625000000000 \\
F(5) &\approx -0.000400000000000 \\
F(6) &\approx -0.000277777777778 \\
F(7) &\approx -0.000204081632653 \\
F(8) &\approx -0.000156250000000 \\
F(9) &\approx -0.000123456790123 \\
F(10) &\approx -0.000100000000000
\end{aligned}
$$

The MATLAB code can be found in the appendix.

Appendix

Listing 1: Code for Least Action

```
addpath('..');
x = 1:1:10;
m=0.1;k=0.01;
y = -k ./ (x.^2);
n = length(y);
y = y';
cvx_setup
cvx_begin
variable F(n)
minimize norm(-F+y,2)
cvx_end
```