DSA3102 Homework 2 A0239929J

1. For which sets of $\alpha \in \mathbb{R}$ are the following functions convex? Explain your answers carefully.

(a) $f(x) = \sin x + \alpha x$ with $\mathbf{dom}\, f = \mathbb{R}$

(b) $f(x) = \sin x + \alpha x^2$ with $\mathbf{dom}\, f = \mathbb{R}$

(c) $f(x_1, x_2) = (5-\alpha)x_1^2 + 10x_1x_2 + x_2^2 + 4\alpha x_1$ with $\mathbf{dom}\, f = \mathbb{R}^2$

(a)

$f(x) = \sin x + \alpha x \quad \mathrm{dom} f = \mathbb{R}$

$f'(x) = \cos x + \alpha$

$f''(x) = -\sin x$

There are no sets of α that will gurantee that $f''(x) > 0$ as $f''(x)$ is independent of α

(b) $f(x) = \sin x + \alpha x^2 \quad \mathrm{dom} f = \mathbb{R}$

$f'(x) = \cos x + 2\alpha x$

$f''(x) = -\sin x + 2\alpha$

For $f(x)$ to be convex, $f''(x) \geq 0$

$\Rightarrow -\sin x + 2\alpha \geq 0$

since $-1 \leq \sin x \leq 1 \Rightarrow -1 \leq -\sin x \leq 1$

for $-\sin x + 2\alpha \geq 0$, $\alpha \geq \frac{1}{2}$

(c) $f(x_1, x_2) = (5-\alpha)x_1^2 + 10x_1x_2 + x_2^2 + 4\alpha x_1$

with $\operatorname{dom} f = \mathbb{R}^2$

$$\nabla f(x,y) = \begin{bmatrix} 2(5-\alpha)x_1 + 10x_2 + 4\alpha \\ 10x_1 + 2x_2 \end{bmatrix}$$

$$\nabla^2 f(x,y) = \begin{bmatrix} 2(5-\alpha) & 10 \\ 10 & 2 \end{bmatrix}$$

By 2nd order optimality, for f to be convex

$2(5-\alpha) > 0 \quad - ①$

and $2(5-\alpha)(2) - 10(10) > 0 \quad - ②$

From ①, $5-\alpha > 0$

$5 > \alpha$

From ② $4(5-\alpha) > 100$

$5-\alpha > 25$

$-20 > \alpha$

Hence the set of α is $\alpha < -20$ in order for f to be convex

2. (Ky Fan's inequality)

Let $x_i \in (0, 1/2]$ and let $\gamma_i \in (0, 1)$ for $i = 1, \ldots, n$ be real numbers satisfying $\sum_{i=1}^{n} \gamma_i = 1$. It is known that either

$$\text{(A)} \qquad \frac{\prod_{i=1}^{n} x_i^{\gamma_i}}{\prod_{i=1}^{n} (1 - x_i)^{\gamma_i}} \le \frac{\sum_{i=1}^{n} \gamma_i x_i}{\sum_{i=1}^{n} \gamma_i (1 - x_i)}$$

or

$$\text{(B)} \qquad \frac{\prod_{i=1}^{n} x_i^{\gamma_i}}{\prod_{i=1}^{n} (1 - x_i)^{\gamma_i}} \ge \frac{\sum_{i=1}^{n} \gamma_i x_i}{\sum_{i=1}^{n} \gamma_i (1 - x_i)}$$

is true. Which is true and why?

Hint: Consider the convexity/concavity properties of the function $f(x) = \ln x - \ln(1 - x)$ with **dom** $f =$ *$[0, 1/2]$ and then apply Jensen's inequality.*

(A) is true.

We will apply Jensen's inequality to the strictly concave function $f(x) = \ln x - \ln(1-x)$ for $\text{dom} f = [0, \frac{1}{2}]$

case 1: If at least 1 x_i is 0, then LHS of (A) is 0 and inequality holds. Equality holds $\Leftrightarrow$ RHS $= 0$ and this is obtained when $\gamma_i x_i = 0$ for $i = 1 \ldots n$

case 2:

Assume $\forall x_i > 0$ and that $\gamma_i > 0$ for $\forall i$

If $x_1 = x_2 = \ldots x_n$, then equality holds

WTS: strict inequality if not all x_i are equal

consider $f(x) = \ln x - \ln(1-x) = \ln\left(\frac{x}{1-x}\right)$

$$f'(x) = \frac{1}{x} + \frac{1}{1-x}$$

$$f''(x) = -\frac{1}{x^2} + \frac{1}{(1-x)^2} < 0 \quad \text{which is concave.}$$

with Jensen's inequality,

$$\sum_{i=1}^{n} \varphi_i f(x_i) \leq f\left(\sum_{i=1}^{n} \varphi_i x_i\right)$$

LHS : $\sum \varphi_i f(x_i) = \sum_{i=1}^{n} \varphi_i \ln\left(\frac{x_i}{1-x_i}\right)$

$$= \ln\left\{\frac{\prod x_i^{\varphi_i}}{\prod (1-x_i)^{\varphi_i}}\right\}$$

RHS : $f\left(\sum_{i=1}^{n} \varphi_i x_i\right) = \ln\left\{\frac{\sum_{i=1}^{n} \varphi_i x_i}{1-\sum_{i=1}^{n} \varphi_i x_i}\right\}$

$$= \ln\left\{\frac{\sum_{i=1}^{n} \varphi_i x_i}{\sum_{i=1}^{n} \varphi_i - \sum_{i=1}^{n} \varphi_i x_i}\right\} \quad \text{since } \sum_{i=1}^{n} \varphi_i = 1$$

$$= \ln\left\{\frac{\sum_{i=1}^{n} \varphi_i x_i}{\sum_{i=1}^{n} \varphi_i (1-x_i)}\right\}$$

Hence

$$\ln\left\{\frac{\prod x_i^{\varphi_i}}{\prod (1-x_i)^{\varphi_i}}\right\} < \ln\left\{\frac{\sum_{i=1}^{n} \varphi_i x_i}{\sum_{i=1}^{n} \varphi_i (1-x_i)}\right\}$$

Taking exponent on both sides gives (A)

3. (Conjugate of the conjugate of a quadratic on bounded interval)
Consider the convex function $f : [0, 1] \to \mathbb{R}$ defined as

$$f(x) = \frac{1}{2}x^2.$$

The domain of f is $\mathbf{dom}\, f = [0, 1]$.

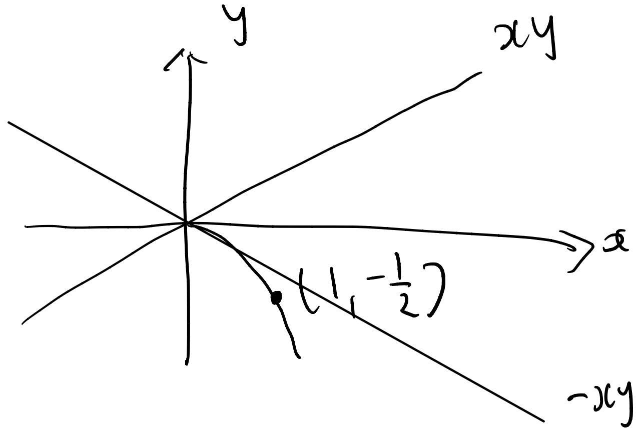
(a) Find the conjugate function $f^*(y)$ specifying various ranges of y carefully.

(b) Now find the conjugate of f^*. Call it f^{**}.

(c) Is $f^{**}(x)$ the same as $f(x)$?

(a) $f^*(y) = \sup_{x \in \mathrm{dom} f} \{x^T y - f(x)\}$

$= \sup_{x \in [0,1]} \{xy - \frac{1}{2}x^2\}$



consider 2 cases

case ①: $y \leq 0$

let $g(x) = xy - \frac{1}{2}x^2$

$g(x)$ maximized when xy is largest and $\frac{1}{2}x^2$ is smallest

since $0 \leq x \leq 1$, xy is largest when $y = 0$ and $x = 0$.

$f^*(y) = 0$

case ②: $y > 0$

let $g(x) = xy - \frac{1}{2}x^2$

$g'(x) = 0 \Rightarrow y - x = 0$

$x^* = y$

$f^*(y) = y(y) - \frac{1}{2}(y)^2 = \frac{1}{2}y^2$

Hence $f^*(y) = \begin{cases} 0 & \text{when } y \leq 0 \\ \frac{1}{2}y^2 & \text{when } y > 0 \end{cases}$

(b) $f^{**}(x) = \sup_{y \in \mathbb{R}} \{ xy - f^*(y) \}$

maximising $xy - \frac{1}{2}y^2 \Leftrightarrow$ maximising xy

Hence we consider

$f^{**}(x) = \sup_{y \in \mathbb{R}} \{ xy - \frac{1}{2}y^2 \}$

case ①: $x < 0$

when $x \to -\infty$ and y is a small -ve value,

$xy - \frac{1}{2}y^2 \to \infty$

$f^{**}(x) = +\infty$ (unbounded above)

case ②: $0 \leq x \leq 1$

From (a), $f^{**}(x) = x(x) - \frac{1}{2}(x)^2$

$= \frac{1}{2}x^2$

case ③: $x > 1$

when $x \to \infty$, $xy - \frac{1}{2}y^2 \to \infty$

$f^{**}(x) = +\infty$ (unbounded above)

$$f^{**}(x) = \begin{cases} \frac{1}{2}x^2 & x \in [0,1] \\ \infty & \text{o/w} \end{cases}$$

1(c) $f^*(y)$ is the same as $f^{**}(x)$ if its domain is the same at $x \in [0,1]$

otherwise the 2 values are not the same

4. (Reformulating a problem into an SOCP)

Consider the following optimization problem:

$$\begin{aligned} \min_x \max_{k=1,\dots,n} & \left|\log(a_k^T x) - \log(b_k)\right| \\ \text{subject to} \quad & x \succeq 0, \end{aligned}$$

where we assume that $b_i > 0$ and $\log(a_i^T x) = -\infty$ when $a_i^T x \leq 0$. Formulate this problem as a second-order cone program (SOCP).

SOCP $\min_{x \in \mathbb{R}^n} f^T x$ s.t $\|A_i x + b_i\| \leq c_i^T x + d_i$, $i \in [m]$

$$\min_{x} \max_{k=1,..n} |\log(a_k^T x) - \log(b_k)|$$

$$= \min_{x} \max_{k=1,..n} \left|\log\left(\frac{a_k^T x}{b_k}\right)\right|$$

$$= \min_{x} \max_{k=1,..n} \log \max\left\{\frac{a_k^T x}{b_k}, \frac{b_k}{a_k^T x}\right\}$$

$$= \min_{x} \max_{k=1,..n} \max\left\{\frac{a_k^T x}{b_k}, \frac{b_k}{a_k^T x}\right\}$$

$$\Rightarrow \min_{x \in \mathbb{R}^n, t \in \mathbb{R}} t \text{ s.t } \max_{k=1,..n} \left\{\frac{a_k^T x}{b_k}, \frac{b_k}{a_k^T x}\right\} \leq t \text{ and } x \geq 0$$

constrain:

$$\frac{a_k^T x}{b_k} \leq t \quad \text{or} \quad \frac{b_k}{a_k^T x} \leq t$$

$$\Rightarrow t - \frac{a_k^T x}{b_k} \geq 0 \quad \text{or} \quad \frac{1}{t} \leq \frac{a_k^T x}{b_k}$$

consider $\frac{1}{t} \leq \frac{a_k^T x}{b_k}$,

$$1 \leq t\left(\frac{a_k^T x}{b_k}\right) \Rightarrow 4 \leq 4t\left(\frac{a_k^T x}{b_k}\right)$$

$$4 + t^2 - 2t\left(\frac{a_k^T x}{b_k}\right) + \left(\frac{a_k^T x}{b_k}\right)^2 \leq t^2 + 2t\frac{a_k^T x}{b_k} + \left(\frac{a_k^T x}{b_k}\right)^2$$

$$\Rightarrow 2^2 + \left(t + \frac{a_k^T x}{b_k}\right)^2 \leq \left(t + \frac{a_k^T x}{b_k}\right)^2$$

$$\sqrt{2^2 + \left(t + \frac{a_k^T x}{b_k}\right)^2} \leq \left(t + \frac{a_k^T x}{b_k}\right)$$

Hence $\min\limits_{(x,t)} t$ s.t $\left\| 2 + \left(t + \frac{a_k^T x}{b_k}\right) \right\| \leq t + \frac{a_k^T x}{b_k}$

$\forall k = 1, 2 \ldots n$ and $x \geq 0$

5. (Reformulating a non-convex problem into a convex problem)

Consider the following optimization problem

$$\begin{aligned} \min_x \ & f_0(x) \\ \text{subject to} \ & f_i(x) \le 0 \quad i = 1, 2, \ldots, m \\ & x \succeq 0 \end{aligned}$$

where each f_i is quadratic, i.e.,

$$f_i(x) = \frac{1}{2} x^T P_i x + q_i^T x + r_i$$

where $P_i \in \mathbf{S}^n$ ($\mathbf{S}^n$ is the set of $n \times n$ symmetric matrices), $q_i \in \mathbb{R}^n$ and $r_i \in \mathbb{R}$ for all $i = 0, 1, \ldots, m$. We note here that the problem may not be convex since we have not restricted P_i to be positive semidefinite.

Suppose that $q_i \preceq 0$ (all the components of the vector q_i are nonpositive) and P_i has nonpositive off-diagonal elements for all $i = 0, 1, \ldots, m$. Reformulate this problem as a convex problem by considering a change of variables from x_i to $y_i = g(x_i)$ for some suitable function g.

Let $y = x_i^2$ since $x \geq 0$ $\rightarrow$ equality constrain

$$f_i(x) = \frac{1}{2} x^T P_i x + q_i^T x + r_i$$

$$= \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} P_{ij} x_i x_j + \sum_{i=1}^{n} q_i x_i + r$$

$$= \frac{1}{2} \sum_{i=1}^{n} P_{ii} x_i^2 + \frac{1}{2} \sum_{i \neq j} P_{ij} x_i x_j + \sum_{i=1}^{n} q_i x_i + r$$

$$= \frac{1}{2} \sum_{i=1}^{n} P_{ii} y_i + \frac{1}{2} \sum_{i \neq j} P_{ij} \sqrt{y_i y_j} + \sum_{i=1}^{n} q_i \sqrt{y_i} + r$$

for us to model a convex problem, the constrains must be convex

$\frac{1}{2} \sum_{i=1}^{n} P_{ii} y_i$ and $\sum_{i=1}^{n} q_i \sqrt{y_i} + r$ are convex

So it suffices to show $\sum_{i \neq j} P_{ij} \sqrt{y_i y_j}$ convex

Set P_{ij} to be in $R^2 \Rightarrow 2 \times 2$ matrix

$$\sum_{i \neq j}^{n} P_{ij} \sqrt{y_i y_j} = P_{12} \sqrt{y_1 y_2} + P_{21} \sqrt{y_2 y_1}$$

$$= (P_{12} + P_{21}) \sqrt{y_1 y_2}$$

$$= X \sqrt{y_1 y_2} \quad \text{where } X = (P_{12} + P_{21})$$

let $h(y_1, y_2) = X \sqrt{y_1 y_2}$

$$h'(y_1, y_2) = \begin{bmatrix} \dfrac{X\sqrt{y_2}}{2\sqrt{y_1}} \\ \dfrac{X\sqrt{y_1}}{2\sqrt{y_2}} \end{bmatrix}$$

$$h''(y_1, y_2) = \begin{bmatrix} -\dfrac{X\sqrt{y_2}}{4(y_1)^{3/2}} & \dfrac{X}{4\sqrt{y_1 y_2}} \\ \dfrac{X}{4\sqrt{y_1 y_2}} & -\dfrac{X\sqrt{y_1}}{4(y_2)^{3/2}} \end{bmatrix}$$

$$\text{Hessian} = \frac{X^2 \sqrt{y_1 y_2}}{16 (y_1 y_2)^{3/2}} - \frac{X^2}{4(y_1 y_2)}$$

$$= \frac{X^2}{16} \left\{ \frac{1}{y_1 y_2} - \frac{1}{y_1 y_2} \right\} = 0 \quad \geq 0$$

Hence hessian is +ve semi definite

thus $\sum_{i \neq j}^{n} P_{ij} \sqrt{y_i y_j}$ is convex

So $f_i(y) = \frac{1}{2}\sum_{i=1}^{n} P_{ii} y_i + \frac{1}{2}\sum_{i \neq j} P_{ij}\sqrt{y_i y_j} + \sum_{i=1}^{n} q_i \sqrt{y_i} + r$

is convex.

we can then formulate the convex opt problem:

$\min_{y} f_0(y)$ s.t $f_i(y) \leq 0$ and

$y = x_i^2$, $x \geq 0$ $\forall i = 1, 2 \ldots m$