CSCI 5254 Homework 2

Tuguluke Abulitibu

September 22, 2020

Chapter 3, Definition of convexity

3.1

(a)

This is just another form of Jensen's inequality

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y)$$

let $\theta = \dfrac{b-x}{b-a}$, then $1-\theta = \dfrac{x-a}{b-a}$, and let $x = a, y = b$ in Jensen's inequality, we will have

$$f(x) \leq \frac{b-x}{b-a}f(a) + \frac{x-a}{b-a}f(b)$$

(b)

From part (a), we will see

$$\frac{f(x)-f(a)}{x-a} \leq \frac{\dfrac{b-x}{b-a}f(a) + \dfrac{x-a}{b-a}f(b) - f(a)}{x-a}$$

$$= \frac{x(f(b)-f(a)) - a(f(b)-f(a))}{(b-a)(x-a)} = \frac{f(b)-f(a)}{b-a}$$

Same would be apply for the second part

$$\frac{f(b)-f(x)}{b-x} \geq \frac{f(b) - f\dfrac{b-x}{b-a}f(a) - \dfrac{x-a}{b-a}f(b)}{b-x}$$

$$= \frac{b(f(b)-f(a)) - x(f(b)-f(a))}{(b-a)(b-x)} = \frac{f(b)-f(a)}{b-a}$$

Combining both, we would have

$$\frac{f(x)-f(a)}{x-a} \leq \frac{f(b)-f(a)}{b-a} \leq \frac{f(b)-f(x)}{b-x}$$

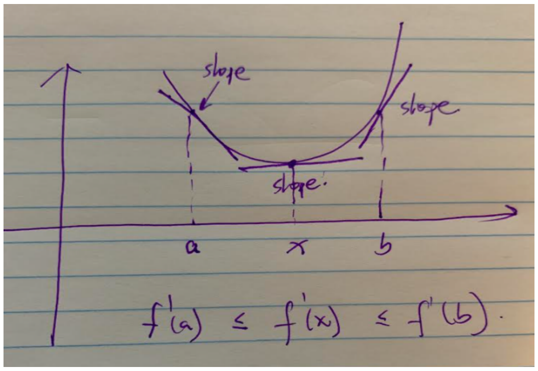

Figure 1: My original draw, based on sign of the slopes

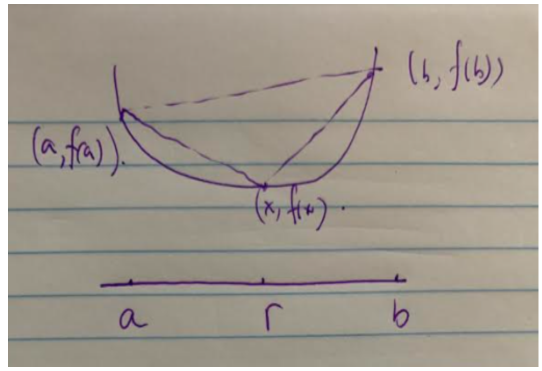

Figure 2: Homework session: a better one with triangle

(c)

Since $a \leq x \leq b$, and from part (b) we have $\dfrac{f(x)-f(a)}{x-a} \leq \dfrac{f(b)-f(a)}{b-a}$, by the definition of the limit, we have

$$f'(a) = \lim_{x \to a} \frac{f(x)-f(a)}{x-a} \leq \frac{f(b)-f(a)}{b-a}$$

similarly:

$$f'(b) = \lim_{x \to b} \frac{f(x)-f(b)}{x-b} \geq \frac{f(b)-f(a)}{b-a}$$

Combining both, we would have

$$f'(a) \leq \frac{f(b)-f(a)}{b-a} \leq f'(b)$$

(d)

From part (c), we have $f'(b) \geq f'(a) \Rightarrow f'(b) - f'(a) \geq 0$, also $b - a\, 0$ by definition of the problem, combine them both, we have $\dfrac{f'(b)-f'(a)}{b-a} \geq 0$. Again we can say this is the definition of second order derivative, hence

$$f''(a) \geq 0,\ f''(b) \geq 0$$

Chapter 3, Examples

3.15

(a)

$$\lim_{\alpha \to 0} u_\alpha(x) = \lim_{\alpha \to 0} \frac{x^\alpha - 1}{\alpha} =^{1} \lim_{\alpha \to 0} \frac{\ln x \times x^\alpha}{1} = \ln x.$$

(b)

Let's look at the first derivative

$$u'_\alpha(x) = \frac{\alpha^2 x^{\alpha-1}}{\alpha} = \alpha x^{\alpha-1}$$

since $0 < \alpha \leq 1$ and $\textbf{dom} u_0 = \mathbf{R}_{++}$, we have $u'_\alpha(x) > 0$, this means the function is monotone increase. Now the second derivative

$$u''_\alpha(x) == \alpha(\alpha-1)x^{\alpha-1} \leq 0$$

[1] L'Hôpital's rule

which means $-u''_\alpha(x) \geq 0$, hence $-u_\alpha$ is convex, this shows that u_α is concave. Finally it is really easy to see $u_\alpha(1) = \dfrac{1-1}{\alpha} = 0$

3.16

We calculate the hessian and see if it is PSD by definition[2]. Let vector $\begin{bmatrix} a & b \end{bmatrix}$ be an arbitrary non-zero vector in this case

(b)

$$\nabla^2 f(x_1, x_2) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

since

$$\begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} b & a \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = 2ab$$

Since it can be any number, hence $f(x_1, x_2)$ is neither convex nor concave.

(c)

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

since

$$\begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} \frac{2}{x_2 x_1^3} & \frac{1}{x_1^2 x_2^2} \\ \frac{1}{x_1^2 x_2^2} & \frac{2}{x_1 x_2^3} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \frac{2a^2}{x_2 x_1^3} + \frac{2ab}{x_1^2 x_2^2} + \frac{2b^2}{x_1 x_2^3} \geq 0 \ \forall a, b \in \mathbf{R}$$

Hence[3] we know $f(x_1, x_2)$ is convex but not concave.

(e)

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

since

$$\begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} \frac{2}{x_2} & \frac{-2x_1}{x_2^2} \\ \frac{-2x_1}{x_2^2} & \frac{2x_1^2}{x_2^2} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \frac{2a^2}{x_2^2} - \frac{4abx_1}{x_2^2} + \frac{2b^2 x_1^2}{x_2^2} \geq 0^4 \ \forall a, b \in \mathbf{R}$$

Hence we know $f(x_1, x_2)$ is convex but not concave.

3.18

Adaptation of Log-determinant on page 74.

[2] M positive semi-definite $\iff$ $x^\mathsf{T} M x \geq 0$ for all $x \in \mathbb{R}^n$

[3] From homework session: the determinant of the middle matrix is $\frac{4}{x_1^2 x_2^2} . - \frac{1}{x_1^2 x_2^2} \geq 0$, which means PSD.

[4] Triangle inequality: $(\frac{\sqrt{2}a}{x_2})^2 + (\frac{\sqrt{2}bx_1}{x_2})^2 \geq \frac{4abx_1}{x_2^2}$

(a)

Approach 1:

$$\begin{aligned} g(t) &= \mathbf{tr}(Z+tV)^{-1} \\ &= \mathbf{tr}[Z^{1/2}(I+tZ^{-1/2}VZ^{-1/2})Z^{1/2}]^{-1} \\ &= \sum_{i=1}^{n}(1+t\lambda_i)^{-1}\cdot \mathbf{tr}Z^{-1} {}^{5} \end{aligned}$$

where $\lambda_1,\ldots,\lambda_n$ are eigenvalues of $Z^{1/2}V^{-1}Z^{1/2}$, therefore we have

$$g'(t) = -\sum_{i=1}^{n}\frac{\lambda_i}{(1+t\lambda_i)^2}\mathbf{tr}(Z^{-1}),\ g''(t) = \sum_{i=1}^{n}\frac{2\lambda_i^2}{(1+t\lambda_i)^3}\mathbf{tr}(Z^{-1}) \text{ where } \mathbf{tr}(Z^{-1})>0$$

Here, with some special t, $g''(t)\geq 0$, we have $f(X)$ as convex. Which is a bit not clear. So I mimic Prof's solution.
Approach 2:

$$\begin{aligned} g(t) &= \mathbf{tr}(Z+tV)^{-1} \\ &= \mathbf{tr}[Z^{-1}(Q(I+t\Sigma)Q^T)^{-1}] \text{ where(through diagonalize) } Z^{-1/2}VZ^{1/2} = Q\Sigma Q^T \\ &= \mathbf{tr}[\overbrace{Q^TZ^{-1}Q}^{A}(I+t\Sigma)^{-1}] = \mathbf{tr}(A(I+t\Sigma)^{-1}) = \sum_i \underbrace{A_{ii}}_{x_1}\underbrace{\frac{1}{1+t\lambda_i}}_{x_2} \end{aligned}$$

Linear functions, hence Convex.

(b)

$$g(t) = (\det(Z+tV))^{1/n} = (\det(Z^{1/2}(I+tZ^{-1/2}Vz^{-1/2})z^{1/2}))^{1/n}$$

$$(\det(Z^{1/2}))^{1/n}\det(I+tZ^{-1/2}VZ^{-1/2}))^{1/n}(\det(Z^{1/2}))^{1/n} = (\prod_{i=1}^{n}(1+t\lambda_i))^{1/n}(\det(Z^1))^{1/n 6}$$

where $\lambda_1,\ldots,\lambda_n$ are eigenvalues of $Z^{1/2}V^{-1}Z^{1/2}$.
From text book page 73: The geometric mean $f(x) = (\prod_{i=1}^{n}(x_i)^{1/n}$ is convex on $\mathbf{dom} f = \mathbf{R}^n_{++}$, and $(\det(Z^1))^{1/n}\geq 0$, hence concave.

3.19

Continuation of Example 3.6 on page 80

(a)

we can rewrite

$$f(x) = \alpha_1x_{[1]}+\alpha_2x_{[2]}+\cdots\alpha_rx_{[r]} = \alpha_r(x_{[1]}+x_{[2]}+\cdots x_{[r]})+(\alpha_{r-1}-\alpha_r)(x_{[1]}+x_{[2]}+\cdots x_{[r-1]})+\cdots+(\alpha_1-\alpha_2)x_{[1]}$$

From the hint we already know that $f(x) = \sum_{i=1}^{k}x_{[i]}$ is convex, and all the coefficient above $(\alpha_{k-1}-\alpha_k)\geq 0$. Linear combination of convex functions with nonnegative coefficients, hence $f(x) = \sum_{i=1}^{r}\alpha_ix_{[i]}$ is convex.

[5] Trace of a matrix is the sum of its eigenvalues
[6] Determinant of A is equal to the product of its eigenvalues

3.22

(b)

Solution is already in the hint.

$$f(x,u,v) = -\sqrt{uv - x^T x} = -\sqrt{u(v - \frac{x^T x}{u})}$$

Let $x_1 = u$ and $x_2 = v - \frac{x^T x}{u}$, by the definition of composition $f(x) = h(g(x))$, in this case:

$$h = \sqrt{x_1 x_2}, \text{ convex and nonincreasing}$$

and

$$g(x) = \begin{cases} g(x_1) = u & \text{concave and convex} \\ g(x_2) = v - \frac{x^T x}{u} & \text{concave since} \frac{x^T x}{u} \text{is convex} \end{cases}$$

by (3.10) on page 84:

- f is convex of h is convex and nonincreasing, and g is concave.

We have $f(x,u,v)$ is convex.

(c)

From part (b) we know that $f(x,u,v) = -\sqrt{uv - x^T x} = -\sqrt{u(v - \frac{x^T x}{u})}$ is convex. We also know from Example 3.13 from page 86 that:

- If g is convex then $-\log(-g(x))$ is convex on $\{x \mid g(x) \le 0\}$

in this case $g = -\sqrt{uv - x^T x} = -\sqrt{u(v - \frac{x^T x}{u})}$, taking the log of $-g$, we have

$$\frac{1}{2}\log(uv - x^T x)$$

hence$-\frac{1}{2}\log(uv - x^T x)$ is convex. Multiplying by a positive constant 2 will not change the convexity.
We showed that $-\log(uv - x^T x)$ is convex.

3.24

(c)

From HW1: 2.15, and from homework session:

$$f(p) = \mathbf{prob}(\alpha \le x \le \beta) = \sum_{k=i}^{j} p_k \text{ where } i = \min\{k \mid \alpha_k \ge \alpha\},\ j = \max\{k \mid \alpha_k \le \beta\}$$

we can easily find the corresponding k s.t. $\alpha_k \ge \alpha$ and $\alpha_k \le \beta$.
Therefore $f(p) = \sum_{k=i}^{j} p_k$ is finite combinations of linear function[7], hence convex, quasiconvex, concave, quasiconcave.

[7]Linear function is 'everything'

(h):Only show that the function is quasiconcave

f is quasiconcave if (1) $\mathbf{dom} f$ is convex, and (2) superlevel sets $C_\alpha = \{x \in \mathbf{dom} f \mid f(x) \geq \alpha\}$ are convex.

$$f(p) = \inf\{\beta - \alpha \mid \mathbf{prob}(alpha \leq x \leq \beta) \geq 0.9\} \geq \gamma^8$$

This is equivalent to

$$\sum_{k=i}^{j} p_k \leq 0.9, \forall i, j \text{ s.t. } \alpha_j - \alpha_i \leq \gamma$$

From part (a), we know we can easily find the corresponding k s.t. $\alpha_k \geq \alpha$ and $\alpha_k \leq \beta$, which satisfies the superlevel set

$$S_r = \{p \mid f(p) \geq \gamma\}$$

hence $\inf\{\beta = \alpha \mid \mathbf{prob}(\alpha \leq x \leq \beta) \geq 0.9\}$ is quasiconcave.

Chapter 3, Conjugate functions

3.26

(a)

The answer is in the hint:

$$\sum_{i=1}^{k} \lambda_i(X) = \sup\{\mathbf{tr}(V^T X V) \mid V \in \mathbf{R}^{n\times k}, V^T V = I\}$$

since

$$\mathbf{tr}(V^T X V) = \mathbf{tr} \begin{bmatrix} (v_1)^T x_{11} v_1 & & \\ & \ddots & \\ & & (v_k)^T x_{ii} v_k \end{bmatrix} = \sum_{i=1}^{k} (v_i)^T x_{ii} v_i$$

Similar to Example 3.10 from page 82, we know that $f(x) = \sum_{i=1}^{k} \lambda_i(X)$ is the pointwise supremum of a family of linear function of X, hence convex.

3.36

(a)

$$f^*(y) = \sup_{x \in \mathbf{R}^n} (y^T x - f(x)) = \sup_{x \in \mathbf{R}^n} (\sum_{i=1}^{n} y_i x_i - \max_{i=1,\dots,n} x_i)$$

consider:

- if $y_i < 0$, then $f^*(y) = \infty$ for some negative x_i
- if $y_i \geq 0$ and $\sum_{i=1}^{n} y_i > 1$, then $f^*(y) = \infty$ for some positive x_i
- if $y_i \geq 0$ and $\sum_{i=1}^{n} y_i < 1$, then $f^*(y) = \infty$ for some negative x_i
- if $y_i \geq 0$ and $\sum_{i=1}^{n} y_i = 1$, then $f^*(y) = \sum_{i=1}^{n} y_i x_i - \max_{i=1,\dots,n} x_i = 0$

this is an idicator function

$$f^*(y) = \begin{cases} 0 & \text{if } \sum_{i=1}^{n} y_i = 1 \\ \infty & \text{otherwise.} \end{cases}$$

[8] γ is the range

(d)

$$f^*(y) = \sup_{x \in \mathbf{R}^n} (y^T x - f(x)) = \sup_{x \in \mathbf{R}^n} (\sum_{i=1}^{n} y_i x_i - x^p)$$

here, $(yx - x^p)' = 0 \Rightarrow x = (\frac{y}{p})^{\frac{1}{p-1}}$, plug it back in, we would have $y(\frac{y}{p})^{\frac{1}{p-1}} - (\frac{y}{p})^{\frac{p}{p-1}}$

when $p > 1$, consider:

- if $y_i < 0$, then $f^*(y) = 0$ for large p
- if $y_i = 0$, then $f^*(y) = 0$
- if $y_i > 0$, then $f^*(y) = y(\frac{y}{p})^{\frac{1}{p-1}} - (\frac{y}{p})^{\frac{p}{p-1}}$

Hence

$$f^*(y) = \begin{cases} 0 & \text{if } y_i \leq 0 \\ y(\frac{y}{p})^{\frac{1}{p-1}} - (\frac{y}{p})^{\frac{p}{p-1}} & \text{otherwise.} \end{cases}$$

when $p < 0$, consider:

- if $y_i < 0$, then $f^*(y) = y(\frac{y}{p})^{\frac{1}{p-1}} - (\frac{y}{p})^{\frac{p}{p-1}}$
- if $y_i = 0$, then $f^*(y) = 0$
- if $y_i > 0$, then $f^*(y) = \infty$, unbounded

Hence

$$f^*(y) = \begin{cases} 0 & \text{if } y_i = 0 \\ y(\frac{y}{p})^{\frac{1}{p-1}} - (\frac{y}{p})^{\frac{p}{p-1}} & \text{if } y_i < 0 \end{cases}$$