98101588 HW3 محمدامین رامی

-1

.1

داریم: $f(x) = \left(\prod_{k=1}^{n} x_k\right)^{\frac{1}{n}}$ حال ماتریس هسین $f(x)$ را حساب می‌کنیم. داریم:

$$\frac{\partial f}{\partial x_i} = \frac{f^n(x)}{x_i} \times \frac{1}{n} \times \frac{f(x)}{f^n(x)} = \frac{f(x)}{n x_i}$$

$$\Longrightarrow \text{if } i \neq j \Longrightarrow \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{f(x)}{n^2} \times \frac{1}{x_i x_j}$$

اگر $i = j$ باشد داریم:

$$\frac{\partial^2 f(x)}{\partial x_i^2} = \frac{\frac{f(x)}{n x_i} n x_i - n f(x)}{n^2 x_i^2} = \frac{f(x)}{n^2} (1-n) \frac{1}{x_i^2}$$

$$\Longrightarrow \nabla^2 f(x) = A - \operatorname{diag}(x^{-2}) \quad \text{where } A_{ij} = \frac{1}{x_i x_j}$$

حال فرض کنید y برداری باشد که $y_i = \frac{1}{x_i}$ آنگاه: $A = y y^T$ حال داریم: $\forall v \in \mathbb{R}^n$:

$$v^T \nabla^2 f(x) v = \left(\sum_{i=1}^{n} y_i v_i\right)^2 - n \sum_{i=1}^{n} y_i^2 v_i^2$$

حال فرض کنید z برداری باشد که از ضرب درایه به درایه از v و y حاصل شده باشد. حال طبق نامساوی کوشی شوارتز داریم:

$$(1^T z) \leq \|1\|^2 \|z\|^2 \Longrightarrow \left(\sum_{i=1}^{n} y_i v_i\right)^2 \leq n \sum_{i=1}^{n} y_i^2 v_i^2$$

پس می‌توان نتیجه گرفت که $v^T \nabla^2 f(x) v \leq 0$ برای هر $v \in \mathbb{R}^n$. پس تابع داده شده مقعر است.

.2

مانند قسمت قبل نشان می‌دهیم که $\nabla^2 f(x) \preceq 0$ است. داریم:

$$\frac{\partial f}{\partial x_i} = p x_i^{p-1} \frac{1}{p} \left(\sum_{k=1}^{n} x_k^p\right)^{\frac{1}{p}-1} = x_i^{p-1} \left(\sum_{k=1}^{n} x_k^p\right)^{\frac{1}{p}-1}$$

$$\Longrightarrow \frac{\partial^2 f}{\partial x_i \partial x_j} = x_i^{p-1} x_j^{p-1} (1-p) \left(\sum_{k=1}^{n} x_k^p\right)^{\frac{1}{p}-2} \quad \text{for } i \neq j$$

$$\text{if } i=j \Longrightarrow \frac{\partial^2 f}{\partial x_i^2} = (1-p)\, x_i^{2(p-1)} \left(\sum_{k=1}^{n} x_k\right)^{\frac{1}{p}-2} - (1-p)\, x_i^{p-2} \left(\sum_{k=1}^{n} x_k^p\right)^{\frac{1}{p}-1}$$

$$= (1-p) \left(\sum_{k=1}^{n} x_k^p\right)^{\frac{1}{p}-2} \left[x_i^{2(p-1)} - x_i^{p-2} \left(\sum_{k=1}^{n} x_k^p\right) \right]$$

حال بردار دلخواه v را در R^n تصور کنید. ثابت می‌کنیم $v^T \nabla^2 f(x)\, v \leq 0$ است. داریم:

$$v^T \nabla^2 f(x)\, v \leq 0 \Longleftrightarrow (1-p)\left(\sum_{k=1}^{n} x_k^p\right)^{\frac{1}{p}-2} \left[\left(\sum_{k=1}^{n} v_k x_k^{p-1}\right)^2 - \left(\sum_{k=1}^{n} x_k^p\right)\left(\sum_{k=1}^{n} v_k^2 x_k^{p-2}\right) \right] \leq 0$$

$$\Longleftrightarrow \left(\sum_{k=1}^{n} v_k x_k^{p-1}\right)^2 - \left(\sum_{k=1}^{n} x_k^p\right)\left(\sum_{k=1}^{n} v_k^2 x_k^{p-2}\right) \leq 0$$

حال اگر y و z را به صورت روبرو تعریف کنیم:

$$z_k = x_k^{\frac{p}{2}} \quad \text{و} \quad y_k = v_k x_k^{\frac{p}{2}-1}$$

باشد، اگر نامساوی کوشی‌شوارتز را برای z و y بنویسیم، نامساوی آخری بدست می‌آید. پس داریم $v^T \nabla^2 f(x)\, v \leq 0$ است.

3

تابع $g(t)$ را به صورت زیر تعریف می‌کنیم:

$$g(t) = f(X + tV) \quad , V \in S$$

حال داریم:

$$g(t) = f(X+tV) = (\det(X+tV))^{\frac{1}{n}} = \left(\det\left(X^{\frac{1}{2}}\left(I + tX^{-\frac{1}{2}}VX^{-\frac{1}{2}}\right)X^{\frac{1}{2}}\right)\right)^{\frac{1}{n}}$$

$$= \left(\det(X)\det\left(I + tX^{-\frac{1}{2}}VX^{-\frac{1}{2}}\right)\right)^{\frac{1}{n}} = (\det(X))^{\frac{1}{n}}\left(\prod_{i=1}^{n}(1+t\lambda_i)\right)^{\frac{1}{n}}$$

که در روابط بالا، λ_i ها مقادیر ویژه ماتریس $X^{-\frac{1}{2}}VX^{-\frac{1}{2}}$ است. $g(t)$ یک تابع مقعر است زیرا تابع $\left(\prod_{i=1}^{n}(1+t\lambda_i)\right)^{\frac{1}{n}}$ میانگین هندسی است و میانگین هندسی تابعی مقعر است. پس $g(t)$ مقعر است پس $f(X)$ نیز مقعر است.

.4

فرض کنید $\tilde{h}(x) = \log(\sum_{i=1}^{n} e^{x_i})$ باشد طبق فرض سوال این تابع محدب است. حال ماتریس A را ماتریسی فرض کنید که سطرهای آن a_i ها باشند و بردار b را برداری که درایه‌های آن b_i ها باشند. آنگاه داریم:

$$h(x) = \log(\sum_{i=1}^{n} e^{a_i^T x + b_i}) = \tilde{h}(Ax+b)$$

پس $h(x)$ هم یک تابع محدب است.

پس تابع $-h(x) = -\log(\sum_{i=1}^{n} e^{a_i x + b})$ یک تابع مقعر است.

همچنین تابع $g(x) = -\log(x)$ نیز یک تابع محدب و نزولی است.

همچنین داریم: $f(x) = g(-h(x))$ پس طبق قضیه ثابت شده سر کلاس، چون از ترکیب یک تابع مقعر و یک تابع محدب نزولی حاصل شده است، محدب است.

.5

تابع $g(x)$ را به صورت روبه رو در نظر بگیرید: $g(x) = \|x\|_p^p = |x_1|^p + |x_2|^p + \cdots + |x_n|^p$

تابع x^p برای $p>1$ محدب است. پس $g(x)$ هم محدب است. حال تابع perspective را برای تابع $g(x)$ در نظر بگیرید، اگر آن را $f(x,t)$ بنامیم داریم:

$$f(x,t) = t\, g(\frac{x}{t}) = t\, \frac{|x_1|^p + |x_2|^p + \cdots + |x_n|^p}{t^p} = \frac{\|x\|_p^p}{t^{p-1}} \quad \mathrm{Dom}_f = \{(x,t) \mid x \in R^n, t>0\}$$

طبق قضیه‌ای که در کلاس بیان شد، $f(x,t)$ نیز تابع محدب است چون $g(x)$ محدب است.

حال تابع داده شده دقیقا تابع perspective $g(x)$ است. پس محدب است.

-2

فرض کنید تابع گفته شده، تابع ثابت نباشد، آنگاه نقطه‌ای مانند x و y در دامنه f وجود دارد به قسمی که: $f(x) > f(y)$

حال نقطه جدید x' را در نظر بگیرید که به صورت روبه‌رو تعریف می‌شود: $x' = \frac{x-(1-\lambda)y}{\lambda}$

که λ عدد دلخواهی بین 0 و 1 است. حال طبق تعریف تحدب داریم:

$$f(\lambda x' + (1-\lambda)y) \le \lambda f(x') + (1-\lambda) f(y)$$

$$\Rightarrow f(x-(1-\lambda)y + (1-\lambda)y) \le \lambda f(x') + (1-\lambda) f(y)$$

$$\Rightarrow f(x) \le \lambda f(x') + (1-\lambda) f(y)$$

$$\Rightarrow f(x') \ge \frac{f(x)-(1-\lambda) f(y)}{\lambda}$$

واضح است که $f(x)-(1-\lambda)f(y)$ عددی مثبت است. حال چون λ دلخواه است، می‌توانیم λ را به صفر میل دهیم، یعنی $\lambda \to 0^+$. آنگاه عبارت $\frac{f(x)-(1-\lambda)f(y)}{\lambda}$ به بی‌نهایت میل می‌کند پس $f(x')$ به بی‌نهایت میل می‌کند که با کراندار بودن تابع f در تناقض است.

-3

داریم:

$$f^*(Y) = \sup_{X \in S^n_{++}} tr(Y^T X) - tr(X^{-1})$$

حال برای ماکسیمم کردن عبارت فوق، کافی است گرادیان بگیریم و مساوی 0 قرار دهیم چون عبارت داخل ماکسیمم مقعراست. داریم:

$$\frac{\partial}{\partial X}\left[tr(Y^T X) - tr(X^{-1})\right] = Y - \nabla f(X) = Y + X^{-2} = 0$$

$$\Rightarrow Y = -X^{-2} \Rightarrow f^*(Y) = tr(-X^{-2}X) - tr(X^{-1})$$

$$\Rightarrow f^*(Y) = -tr(X^{-1}) - tr(X^{-1}) = -2\,tr(X^{-1}) = -2\,tr(-Y)^{\frac{1}{2}}$$

طبق تعریف داریم:

$$f(\lambda x+(1-\lambda)y) \le \lambda f(x)+(1-\lambda)f(y)$$

$$\Rightarrow f^2(\lambda x+(1-\lambda)y) \le [\lambda f(x)+(1-\lambda)f(y)]^2 \quad \mathrm{I}$$

$$g(\lambda x+(1-\lambda)y) \ge \lambda g(x)+(1-\lambda)g(y)$$

چون g مثبت است داریم:

$$\frac{1}{g(\lambda x+(1-\lambda)y)} \le \frac{1}{\lambda g(x)+(1-\lambda)g(y)} \quad \mathrm{II}$$

چون هر دوی f و g مثبت هستند، از ضرب I و II به رابطه زیر می‌رسیم:

$$\frac{f^2(\lambda x+(1-\lambda)y)}{g(\lambda x+(1-\lambda)y)} \le \frac{[\lambda f(x)+(1-\lambda)f(y)]^2}{\lambda g(x)+(1-\lambda)g(y)}$$

حال کافی است نشان دهیم رابطه زیر درست است:

$$\lambda\frac{f^2(x)}{g(x)}+\frac{(1-\lambda)f^2(y)}{g(y)} \ge \frac{[\lambda f(x)+(1-\lambda)f(y)]^2}{\lambda g(x)+(1-\lambda)g(y)} \iff$$

$$\frac{\lambda^2 f(x)}{\lambda g(x)}+\frac{(1-\lambda)^2 f^2(y)}{(1-\lambda)g(y)}-\frac{\lambda^2 f^2(x)+(1-\lambda)f^2(y)+2\lambda(1-\lambda)f(x)f(y)}{\lambda g(x)+(1-\lambda)g(y)} \ge 0$$

$$\iff \lambda^3(1-\lambda)f^2(x)g(x)g(y)+\lambda^2(1-\lambda)^2f^2(x)g^2(y)+\lambda(1-\lambda)^3f^2(y)g(x)g(y)$$
$$+\lambda^2(1-\lambda)^2f^2(y)g(x)-\lambda^3(1-\lambda)f^2(x)g(x)g(y)-\lambda(1-\lambda)^3f^2(y)g(y)g(x)$$
$$-2\lambda^2(1-\lambda)^2f(x)f(y)g(x)g(y)$$

$$= \lambda^2(1-\lambda)^2 f^2(x) g^2(y) + \lambda^2(1-\lambda)^2 f^2(y) g^2(x) - 2\lambda^2(1-\lambda)^2 f(x) f(y) g(x) g(y) \geq 0$$

$$\Longleftrightarrow \lambda(1-\lambda)\left(f(x)g(y) - f(y)g(x)\right)^2 \geq 0 \Longleftrightarrow \text{True}$$

پس نامساوی زیر برقرار است:

$$\lambda \frac{f^2(x)}{g(x)} + (1-\lambda)\frac{f^2(y)}{g(y)} \geq \frac{f^2(\lambda x + (1-\lambda)y)}{g(\lambda x + (1-\lambda)y)}$$

پس از مقایسه نامساوی بالا و I نتیجه می‌گیریم که $\frac{f^2}{g}$ محدب است.

-5

داریم:

$$f^*(Y) = \sup_{X \in S_{++}^n} tr(Y^T X) - f(X)$$

حال برای اینکه عبارت داخل sup، سوپریمم داشته باشد باید داشته باشیم: $Y \in -S_+^n$، زیرا تابع $-f(X)$ مقعر است و اگر Y نیز منفی معین باشد، عبارت داخل sup مقعر می‌شود و سوپریمم خواهد داشت. پس $\text{Dom} f^* = -S_+^n$. حال $f^*(Y)$ را حساب می‌کنیم:

$$\frac{\partial}{\partial X} tr(Y^T X) - f(X) = Y - \nabla f(X) = Y - (-X^{-1}) = Y + X^{-1} \Rightarrow Y = -X^{-1}$$

$$\Rightarrow f^*(Y) = tr(-Y^{-1}X) + \log(\det(-Y^{-1})) = -tr(I) + \log(\det(-Y^{-1}))$$

$$= -n - \log(\det(-Y)) \Longrightarrow \boxed{f^*(Y) = -n - \log(\det(-Y))}$$

-6

.1

فرض کنید $g(x) = \frac{1}{h(x)}$ حال داریم:

$$g(x) \text{ is concave} \iff g(x) \leq g(y) - g'(y)(y-x) \iff g(x) \leq g(y) + \frac{h'(y)}{h^2(y)}(y-x)$$

$$\iff \frac{1}{h(x)} \leq \frac{1}{h(y)} + \frac{h'(y)}{h^2(y)}(y-x) \iff \frac{h^2(y)}{h(x)} \leq h(y) + h'(y-x)$$

2. فرض $g(x) = \frac{1}{h(x)}$ باشد. حال اگر رابطه گفته شده برقرار باشد داریم:

$$h(y) - h'(y)(x-y) = \frac{h^2(y)}{h(x)} \Rightarrow \frac{1}{h(y)} - \frac{h'(y)}{h^2(y)}(x-y) = \frac{1}{h(x)}$$

$$\Rightarrow g(y) - g'(y)(x-y) = g(x) \Rightarrow g'(y) = \frac{g(y)-g(x)}{y-x} \quad \forall x,y \in I$$

$$\Rightarrow g(y) \text{ is affine}$$

برقرار بودن $h(y) + h'(y)(y-x) = \frac{h^2(y)}{h(x)}$ معادل برقرار بودن $g'(y) = \frac{g(y)-g(x)}{y-x}$ است که رابطه دوم به وضوح برای توابع افاین برقرار است. پس هر دو طرف را ثابت کردیم.

3. فرض کنید $g(x) = \frac{1}{h(x)}$ باشد. داریم:

$$g(x) \text{ is concave} \iff g''(x) \leq 0 \iff \frac{d^2}{dx^2}\left(\frac{1}{h(x)}\right) \leq 0$$

$$\iff \frac{-h''(x)h^2(x) + 2h(x)h'^2(x)}{h^4(x)} \leq 0 \iff 2h(x)h'^2(x) \leq h''(x)h^2(x)$$

$$\iff 2h'^2(x) \leq h(x)h''(x)$$

4.

اگر تابع $\frac{1}{h(x)}$ مقعر باشد، طبق تمرین 3 رابطه زیر برقرار است:

$$h(x)h''(x) \ge 2h'^2(x) \Longrightarrow h''(x) \ge \frac{2h'^2(x)}{h(x)} \ge 0 \Longrightarrow h(x) \text{ is concave}$$

برای مثال نقض نیز فرض کنید $h:(0,\infty)\to(0,\infty)$ باشد و $h(x)=x^2$

واضح است که $h(x)$ محدب است. اما $\frac{1}{h(x)}=\frac{1}{x^2}$ نیز محدب است و مقعر نیست.

7-

داریم:

$$C(s)=\sum_{n=1}^{\infty}\frac{1}{n^{s}} \Longrightarrow c'(s)=-s\sum_{n=1}^{\infty}\frac{1}{n^{s+1}}$$

$$\Longrightarrow c'(s)=-s\,c(s+1) \Longrightarrow c''(s)=-c(s+1)+s^2c(s+2)$$

$$\Longrightarrow c''(s)\ge 0 \Longleftrightarrow s^2\ge\frac{c(s+1)}{c(s+2)} \Longrightarrow \text{if } s\in(1,\infty)$$

$$\Longrightarrow c''(s)\ge 0 \Longrightarrow c(s) \text{ is convex} \Longrightarrow \ln(c(s)) \text{ is convex}$$