به نام خدا

علی قاسم زاده ۴۰۱۱۰۶۳۳۹ تمرین سوم بهینه سازی محدب

۱- می خواهیم برای تابع محدب f روی Polyhedron: $P=\text{conv}\{v_1 \dots v_k\}$ نشان دهیم که:

$$\sup_{x\in P} f(x) = \max_{i=1\dots k} f(v_i)$$

$\text{conv}\{v_1 \dots v_k\}$ عبارت است از تمام $\sum_{i=1}^{k}\lambda_i v_i$ به شرطی که $\forall i: \lambda_i \geq 0$, $\sum_{i=1}^{k}\lambda_i = 1$.

حال فرض کنیم اتفاق بالا نیفتد یعنی یک v ای به صورت $\sum_{i=1}^{k}\lambda_i v_i$ وجود داشته باشد که

$f(v) > f(v_i)$: $\forall i$ پس داریم که ،

$$\forall j: f\left(\sum_{i=1}^{k}\lambda_i v_i\right) > f(v_j)$$

درصورتی که طبق نامساوی Jenson داریم که ،

$$\overbrace{f\left(\sum_{i=1}^{k}\lambda_i v_i\right)}^{f(v)} \leq \sum_{i=1}^{k}\lambda_i f(v_i)$$

پس حداقل یکی از $f(v_j)$ ها هستند که از $f(v)$ بزرگتر مساوی هستند ← ※

زیرا فرض کرده بودیم که $f(v)$ از تمام $f(v_j)$ ها بزرگتر است ،

⇐ حکم برقرار است یعنی : $\sup_{x\in P} f(x) = \max_{i=1\dots k} f(v_i)$ برقرار است .

حالا عبارت قوی تر را اثبات می کنیم: یعنی اگر تابع f محدب باشد روی یک مجموعه ی بسته ی محدب، آنگاه ماکزیمم آن در extreme point ها اتفاق می افتد

اگر نقاط وسط مجموعه یا به اصطلاح interior points ، ماکزیمم داشته باشیم که همانند بالا می شود، زیرا نقاط وسط مجموعه ی convex را می توان به صورت جمع یک نقطه ی مرزی و یک نقطه ی دیگر مجموعه می توان نوشت و حتی می توان به صورت ضریبی از تعدادی از نقاط وسطی نوشت و به تناقض می خوریم.
می توانیم یک نقطه ی مرزی را بگیریم و آنرا به نقطه ی max وصل کنیم و امتداد دهیم تا به طرف دیگر مجموعه ی محدب

Scanned with CamScanner

را قطع کند ← نقطه‌ی max را می‌توانیم به صورت ترکیب convex ای از این دو نقطه بنویسیم.

← حداقل یکی از این نقاط مرزی f اش از f نقطه max بزرگتر مساوی است ←

ماکزیمم حتماً در یکی از نقاط مرزی یا extreme point اتفاق می‌افتد.

۳.۱۷

۲- می‌خواهیم نشان دهیم که به ازای $p<1$ و $p\neq 0$، تابع $f(x)=\left(\sum_{i=1}^{n} x_i^p\right)^{1/p}$ یک تابع concave است، مشتق دوم این تابع را حساب می‌کنیم (ابتدا مشتق اول و سپس دوم).

$$\frac{\partial f(x)}{\partial x_j} = \frac{\partial}{\partial x_j}\Big(\underbrace{\sum_{i=1}^{n} x_i^p}_{u}\Big)^{1/p} = \frac{\partial x_j^p}{\partial x_j}\left(\frac{\partial u^{1/p}}{\partial u}\right) = p x_j^{p-1}\frac{1}{p}\left(\sum_{i=1}^{n} x_i^p\right)^{\frac{1}{p}-1}$$

$$= x_j^{p-1}\left(\sum_{i=1}^{n} x_i^p\right)^{\frac{1}{p}-1} = x_j^{p-1}\left(\left(\sum_{i=1}^{n} x_i^p\right)^{1/p}\right)^{1-p} = \left(\frac{\overbrace{\left(\sum_{i=1}^{n} x_i^p\right)^{1/p}}^{f(x)}}{x_j}\right)^{1-p}$$

$$\rightsquigarrow \boxed{\frac{\partial f(x)}{\partial x_j} = \left(\frac{f(x)}{x_j}\right)^{1-p}}$$

حالا داریم که:

$$\frac{\partial^2 f(x)}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_j}\left(\frac{f(x)}{x_i}\right)^{1-p} = \left(\frac{1}{x_i}\right)^{1-p}\frac{\partial f(x)^{1-p}}{\partial x_j} = \left(\frac{1}{x_i}\right)^{1-p}(1-p)\, f(x)^{-p}\left(\frac{f(x)}{x_j}\right)^{1-p} =$$

$$\frac{(1-p)}{(x_i x_j)^{1-p}} f(x)^{1-2p} = \frac{1-p}{f(x)}\,\frac{f(x)^{2-2p}}{(x_i x_j)^{1-p}} = \frac{1-p}{f(x)}\left(\frac{f(x)^2}{x_i x_j}\right)^{1-p}$$

$$\Rightarrow \boxed{\frac{\partial f(x)}{\partial x_i \partial x_i} = \frac{1-p}{f(x)}\left(\frac{f(x)^2}{x_i x_j}\right)^{1-p}}$$

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

$$+\left(\frac{1}{x_i}\right)^{1-p}(1-p)\, f(x)^{-p}\left(\frac{f(x)}{x_i}\right)^{1-p} =$$

Scanned with CamScanner

اگر عبارت حاصل از $\partial^2 f(x) / \partial x_i \partial x_i$ را بنویسیم داریم که،

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

و داریم که

$$\frac{\partial^2 f(x)}{\partial x_i \partial x_j} = \frac{1-p}{f(x)}\left(\frac{f(x)^2}{x_i x_j}\right)^{1-p}$$

حالا داریم که به ازای هر y دلخواه،

$$y^T \nabla^2 f(x)\, y = \sum_{i,j=1}^{n} y_i \left(\nabla^2 f(x)\right)_{ij} y_j =$$

$$\frac{(1-p)}{f(x)}\underbrace{\left[\sum_{\substack{i,j=1\\ i\neq j}}^{n} y_i \left(\frac{f(x)^2}{x_i x_j}\right)^{1-p} y_j + \sum_{i=1}^{n} y_i \left(\frac{f(x)^2}{x_i x_i}\right)^{1-p} y_i\right]} - (1-p)\sum_{i=1}^{n} \frac{y_i}{x_i}\left(\frac{f(x)}{x_i}\right)^{1-p} y_i$$

$$= \frac{1-p}{f(x)}\left(\sum_{i,j=1}^{n} y_i \left(\frac{f(x)^2}{x_i x_j}\right)^{1-p} y_j - \sum_{i=1}^{n} \frac{y_i^2 f(x)^{2-p}}{x_i^{2-p}}\right) =$$

$p<1$

مثبت

و f توانی از جمع یه سری عدد مثبت است

$$\frac{1-p}{f(x)}\left(\left(\sum_{i=1}^{n} \frac{y_i f(x)^{1-p}}{x_i^{1-p}}\right)^2 - \sum_{i=1}^{n} \frac{y_i^2 f(x)^{2-p}}{x_i^{2-p}}\right) \le 0$$

باید نشان دهیم،

باید نشان دهیم $\longrightarrow$

$$\left(\sum_{i=1}^{n} \frac{y_i f(x)^{1-p}}{x_i^{1-p}}\right)^2 \le \sum_{i=1}^{n} \frac{y_i^2 f(x)^{2-p}}{x_i^{2-p}} \Rightarrow$$

اگر $\underbrace{y_i \left(\frac{f(x)}{x_i}\right)^{1-p/2}}_{v_i}$ ها را درایه های یک بردار در نظر بگیریم،

خواهیم داشت که $\sum_{i=1}^{n} \frac{y_i^2 f(x)^{2-p}}{x_i^{2-p}} = v^T v = \|v\|_2^2$

از طرفی اگر v_iها را از عبارت سمت راستی هم برداریم، خواهیم داشت که،

$$\frac{y_i f(x)^{1-p}}{x_i^{1-p}} = y_i \left(\frac{f(x)}{x_i}\right)^{1-p/2} \underbrace{\left(\frac{f(x)}{x_i}\right)^{-p/2}}_{u_i \text{ ها}}$$

همچنین داریم که

$$u^T u = \frac{\sum_{i=1}^{n} x_i^{p}}{f(x)^{p}} = 1$$

حالا داریم که:

$$\left(\sum_{i=1}^{n} y_i \frac{f(x)^{1-p}}{x_i^{1-p}}\right)^2 = (u^T v)^2 , \quad \sum_{i=1}^{n} \frac{y_i^2 f(x)^{2-p}}{x_i^{2-p}} = (u^T u)(v^T v) = \|u\|_2^2 \|v\|_2^2$$

$$\Longrightarrow \left(\sum y_i \frac{f(x)^{1-p}}{x_i^{1-p}}\right)^2 \le \sum_{i=1}^{n} \frac{y_i^2 f(x)^{2-p}}{x_i^{2-p}} \longleftrightarrow (u^T v)^2 \le \|u\|_2^2 \|v\|_2^2$$

← این عبارت هم طبق نامساوی کوشی شوارتز برقرار است.

← حکم برقرار است. (روابط همگی برگشت پذیر بودند.)

← تابع f روی دامنه R^n_{++}، concave است.

۳.۱۸

a) باید نشان دهیم که: $f(X) = tr(X^{-1})$ تابعی محدب است روی $dom f = S^n_{++}$:

طبق راهنمای سوال باید مشابه اثبات concavity برای log det پیش برویم:

$$g(t) = f(\underbrace{X+tV}_{\in dom f}) , \quad g(t) \text{ is convex if and only if } f(x) \text{ is convex}$$

← دامنه f ماتریس های متقارن مثبت معین هستند، پس $X+tV \succ 0$، همچنین $X \succ 0$

حالا داریم که،

$$tr\left((X+tV)^{-1}\right) = tr\left(X^{-1/2}\left(I+tX^{-1/2}VX^{-1/2}\right)^{-1}X^{-1/2}\right) = tr\left(X^{-1}\left(I+tX^{-1/2}VX^{-1/2}\right)^{-1}\right)$$

از طرفی $V \in S^n$ است همچنین $X^{-1/2}$ هم عضو S^n_{++} هست ← $(X^{-1/2}VX^{-1/2})^T = X^{-1/2}VX^{-1/2}$

می‌توان نوشت ← $X^{-1/2}VX^{-1/2} = Q\Lambda Q^T$,

$$tr\left(X^{-1}(I+tQ\Lambda Q^T)^{-1}\right) = tr\left(X^{-1}(QQ^T+tQ\Lambda Q^T)^{-1}\right) = tr\left(X^{-1}Q(I+t\Lambda)^{-1}Q^T\right)$$

$$= tr\left(Q^T X^{-1} Q (I+t\Lambda)^{-1}\right) = \sum_{i=1}^{n} (Q^T X^{-1} Q)_{ii} \frac{1}{1+t\lambda_i}$$

تابع $\frac{1}{1+t\lambda_i}$ نسبت به t، convex است همچنین $(Q^TX^{-1}Q)_{ii}$ یک ضریب

مستقل از t است ← به ازای هر i، $(Q^TX^{-1}Q)_{ii}\frac{1}{1+t\lambda_i}$ تابعی convex است نسبت به t

← جمع آنها هم نسبت به t محدب است ← $\boxed{\sum_{i=1}^{n}(Q^TX^{-1}Q)_{ii}\frac{1}{1+t\lambda_i}}$ تابعی محدب است.

b) می‌خواهیم نشان دهیم که $f(X)=\det(X)^{1/n}$ روی $\text{dom} f = S^n_{++}$ تابعی concave است،

مطابق بخش قبل عمل می‌کنیم.

$f(X)$ is concave iff $g(t)=f(X+tV)$ is concave

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

$$\det(X^{1/2})^{1/n}\det(I+tX^{-1/2}VX^{-1/2})^{1/n}\det(X^{1/2})^{1/n}=\det(X)^{1/n}\det(I+\underbrace{tX^{-1/2}VX^{-1/2}}_{Q\Lambda Q^T})^{1/n}$$

مطابق قبل $Q\Lambda Q^T$

$$=\det(X)^{1/n}\det\left(Q(I+t\Lambda)Q^T\right)^{1/n}=\det(X)^{1/n}\cancel{\det(Q)^{1/n}}\det(I+t\Lambda)^{1/n}\cancel{\det(Q^T)^{1/n}}$$

$$=\det(X)^{1/n}\det(I+t\Lambda)^{1/n}=\det(X)^{1/n}\boxed{\left[\prod_{i=1}^{n}(1+t\lambda_i)\right]^{1/n}}$$

چون $X\in S^n_{++}$ ← مثبت است.

Geometric mean است ← concave است نسبت به t، و $\det(X)^{1/n}$ هم به t ربطی ندارد، و مثبت است ←

$g(t)$ تابعی concave است.

اثبات اینکه $\left[\prod_{i=1}^{n}(1+t\lambda_i)\right]^{1/n}$ تابعی concave است، با اثبات نشان می‌دهیم که $f(x)=\prod_{i=1}^{n}x_i^{1/n}$ تابعی concave است،

$$\frac{\partial f}{\partial x_i}=\left(\prod_{j\neq i}^{n}x_j^{1/n}\right)\underbrace{\frac{\partial x_i^{1/n}}{\partial x_i}}_{\frac{1}{n}\times\frac{1}{x_i}x_i^{1/n}}=\left(\prod_{j=1}^{n}x_j^{1/n}\right)\frac{1}{n}x_i^{-1}\rightarrow\frac{\partial f}{\partial x_i}=\frac{1}{n}x_i^{-1}f(x)$$

Scanned with CamScanner

$$\xrightarrow{i\neq j} \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{1}{n} x_i^{-1} \frac{\partial f}{\partial x_j} = \frac{1}{n^2} x_i^{-1} x_j^{-1} f(x) \longrightarrow \frac{\partial^2 f(x)}{\partial x_i \partial x_j} = \frac{1}{n^2} x_i^{-1} x_j^{-1} f(x)$$

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

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

$$\rightsquigarrow v^T \nabla^2 f(x)\, v = \sum_{i,j} v_i \left(\nabla^2 f(x)\right)_{ij} v_j = \sum_{i,j} \frac{v_i v_j}{n^2} f(x) \frac{1}{x_i x_j} - \frac{1}{n} \sum_{i=1}^{n} \frac{v_i v_i f(x)}{x_i x_i}$$

$$= \frac{f(x)}{n^2} \overbrace{\left(\sum_{i,j} \frac{v_i v_j}{x_i x_j}\right)}^{\left(\sum_i \frac{v_i}{x_i}\right)^2} - \frac{f(x)}{n} \sum_{i=1}^{n} \frac{v_i^2}{x_i^2}$$

می خواهیم نشان دهیم $f(x)$ concave است پس باید نشان دهیم $\nabla^2 f(x) \preceq 0$ یعنی به ازای هر v
باید داشته باشیم $v^T \nabla^2 f(x) v \leq 0$ ← باید نشان دهیم

$$\frac{f(x)}{n^2}\left(\sum_{i=1}^{n} \frac{v_i}{x_i}\right)^2 - \frac{f(x)}{n} \sum_{i=1}^{n} \frac{v_i^2}{x_i^2} \leq 0 \longleftrightarrow \underbrace{\left(\sum_{i=1}^{n} \frac{v_i}{x_i}\right)^2 \leq n \sum_{i=1}^{n} \frac{v_i^2}{x_i^2}}$$

این رابطه هم طبق نامساوی کوشی برقرار است ← حکم برقرار است ←

$f(x) = \prod_{i=1}^{n} x_i^{1/n}$ تابعی concave است حال تابع $\prod_{i=1}^{n} (1+t\lambda_i)^{1/n}$

ترکیب f با تابع affine است، $X = \underbrace{[1+t\lambda_1, \dots, 1+t\lambda_n]}$

$$X = \vec{1} + t\,[\lambda_1, \dots, \lambda_n]$$

← این تابع هم concave است.

۳- ابتدا نشان می‌دهیم که $f(X,t) = nt\log t - t\log\det X$ به ازای $\text{dom} f = S_{++}^n \times R_{++}$ روی (X,t) تابعی محدب است:

می‌دانیم که t perspective، convexity و concavity را حفظ می‌کند و داریم که:

$$h(X) = -\log\det(X)$$

h تابعی convex است زیرا $\log\det X$ تابعی concave است ←

$L(X,t) = t\,h(X/t)$ is convex too →

$-t\log\det(X/t)$ is convex too →

$-t\log\det(X)/t^n$ is convex too →

$-t(\log\det(X) - n\log t)$ is convex →

$$\boxed{nt\log t - t\log\det(X) \text{ is convex on } (X,t)}$$

می‌دانیم که به ازای هر $t>0$، $L(X,t)$ تابعی محدب است، L همان f است ←

به ازای $t>0$، $X>0$، $f(X,t)$ تابعی محدب است.

از آنجایی که $X \in S_{++}^n$ ← $tr(X) > 0$ ← اگر بجای t، $tr(X)$ را قرار دهیم باز هم تابع حاصل باید convex باشد ←

$$g(X) = n\underbrace{(tr(X))}_{\sum_{i=1}^n \lambda_i}\log\underbrace{(tr(X))}_{\sum_{i=1}^n \lambda_i} - \underbrace{(tr(X))}_{\sum_{i=1}^n \lambda_i}\underbrace{(\log\det(X))}_{\sum_{i=1}^n \log\lambda_i}$$

تابعی محدب است.

$$g(X) = \left(\sum_{i=1}^n \lambda_i\right)\left(n\log\sum_{i=1}^n \lambda_i - \sum_{i=1}^n \log\lambda_i\right)$$ is a convex function

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

تابع $tr(v^T x v)$ نسبت به X خطی است ← محدب است، حالا داریم pointwise ، از تابع مان supremum می‌گیریم ← حاصل هر تابعی محدب است ← $\sum_{i=1}^{k}\lambda_i(x)$ تابعی محدب است.

b)

$$\left(\prod_{i=n-k+1}^{n}\lambda_i(x)\right)^{1/k} = \frac{1}{k}\inf\{tr(v^T x v) \mid v\in R^{n\times k}, \det v^T v = 1\}$$

تابع $tr(v^T x v)$ نسبت به X خطی است ← concave است ، حالا داریم pointwise ، از تابع مان infimum می‌گیریم ← حاصل هر تابعی concave است ← $\left(\prod_{i=n-k+1}^{n}\lambda_i(x)\right)^{1/k}$ تابعی concave است.

c)

$$\prod_{i=n-k+1}^{n} \lambda_i(x) = \inf\left\{\prod_{i=1}^{k}(v^T x v)_{ii} \mid v\in R^{n\times k}, v^T v = I\right\}$$

$$\xrightarrow{\log} \sum_{i=n-k+1}^{n}\log\lambda_i(x) = \inf\left\{\overset{\sum_{i=1}^{k}\log(v^T x v)_{ii}}{\cancel{\log\prod_{i=1}^{k}(v^T x v)_{ii}}} \mid v\in R^{n\times k}, v^T v = I\right\}$$

$\sum_{i=1}^{k}\log(v^T x v)_{ii}$ جمع یه سری تابع concave است ← خودش هم concave است

← داریم از تابع مان pointwise ، infimum می‌گیریم ← حاصل هر تابعی محدب است

← $\sum_{i=n-k+1}^{n}\log\lambda_i(x)$ تابعی concave است.

$\Rightarrow$

۶- a) می‌خواهیم نشان دهیم به ازای $Y, Z \in S^n_{++}$، $0 \prec Y \preceq Z$، برقرار است، $\det Y \le \det Z$

ابتدا نشان می‌دهیم که به ازای $A \in S^n_+$ (ماتریس A PSD) داریم $\det(A+I) \ge 1$،

$$\det(A+I) = \prod_{i=1}^{n} (1+A_{ii})$$

از آنجایی که $A \in S^n_+$ (A، PSD است)، تمام عناصر روی قطر اصلی اصلی باید مثبت باشند. چرا که با ضرب e_i از دو طرف عناصر روی قطر اصلی بدست می‌آید و باید داشته باشیم $e_i^T A e_i \ge 0 \rightarrow A_{ii} \ge 0$.

حالا در نتیجه $1+A_{ii} \ge 1 \rightarrow \prod_{i=1}^{n} (1+A_{ii}) \ge 1$.

حالا داریم که، از آنجایی که Y مثبت معین و متقارن است، $Y = RR$ ، $R \in S^n_{++}$.

$$\det(X) = \det(R R^{-1} X R^{-1} R) = \det(R)\det(R^{-1} X R^{-1})\det(R) =$$

$$\det(R^2)\det(R^{-1} X R^{-1}) = \det(Y)\det(R^{-1} X R^{-1}) = \det(Y)\det(R^{-1} X R^{-1} - I + I)$$

$$= \det(Y)\det(R^{-1} X R^{-1} - R^{-1} Y R^{-1} + I) = \det(Y)\det(\overset{PD}{R^{-1}}\underbrace{(X-Y)}_{PSD}\overset{PD}{R^{-1}} + I) \ge \det(Y)$$

PSD

به ازای هر بردار Y.

$$\underbrace{Y^T R^{-1}}_{z^T}(X-Y)\underbrace{R^{-1} Y}_{z} \rightarrow z^T (X-Y) z \ge 0 \rightarrow R^{-1}(X-Y)R^{-1} \text{ is PSD}$$

$$\Rightarrow \det(X) \ge \det(Y)$$

(b

$$\det Y^{-1} = \frac{1}{\det Y} \longrightarrow \log \det Y^{-1} = \log \frac{1}{\det Y} = -\log \det Y$$

← مینیمم کردن $\log \det Y^{-1}$ معادل ماکزیمم کردن $\log \det Y$ است.

حالا باید با توجه به قیود مسئله $\log \det Y$ را ماکزیمم کنیم.

$$X \in S^n_{++} \longrightarrow X_{11}, X_{22} > 0 \;\text{و}\; S > 0$$

$$\begin{bmatrix} Y & 0 \\ 0 & 0 \end{bmatrix} \preceq \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}$$

$$\longrightarrow \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \preceq \begin{bmatrix} X_{11}-Y & X_{12} \\ X_{21} & X_{22} \end{bmatrix}$$

حالا برای بلوک 11، مکمل شور را می‌نویسیم ← داریم که

$$X_{11} - Y - X_{12} X_{22}^{-1} X_{21} \succeq 0$$

این عبارت باید بزرگتر مساوی صفر باشد یا به عبارتی ماتریس حاصل باید PSD باشد.

←

باید داشته باشیم: $Y \preceq X_{11} - X_{12} X_{22}^{-1} X_{21} \rightarrow$

از سوال قبل داریم که: $\det(Y) \leq \det(X_{11} - X_{12} X_{22}^{-1} X_{21})$

← برای ماکزیمم کردن $\log \det Y$ ما که معادل ماکزیمم کردن $\det Y$ و مینیمم کردن $\log \det Y^{-1}$ است، کافی است Y را قرار دهیم، $X_{11} - X_{12} X_{22}^{-1} X_{21}$. پس جواب مسئله ما $X_{11} - X_{12} X_{22}^{-1} X_{21}$ است.

C) باید ثابت کنیم که $f(X) = \log\det(X_{11} - X_{12}X_{22}^{-1}X_{21})$ تابعی محدب است.

می دانیم که $\log\det Y$ تابعی concave است $\leftarrow$ $\log\det Y^{-1}$ تابعی convex است.

حالا تعریف می کنیم،

$$G(X,Y) = \log\det Y^{-1} \quad , \quad \operatorname{dom} G = \left\{ (X,Y) \;\middle|\; \begin{matrix} Y \succ 0 , \\ X \succ 0 \end{matrix} \; \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} \succeq \begin{bmatrix} Y & 0 \\ 0 & 0 \end{bmatrix} \right\}$$

دامنه G، مجموعه ای محدب است زیرا:

$$\begin{matrix} (X^1,Y^1) \in \operatorname{dom} G \\ (X^2,Y^2) \in \operatorname{dom} G \end{matrix} \longrightarrow \begin{matrix} Y^1 \succ 0 \\ Y^2 \succ 0 \end{matrix} \quad \begin{matrix} X^1 \succ 0 \\ X^2 \succ 0 \end{matrix} \quad \begin{matrix} \begin{bmatrix} Y^1 & 0 \\ 0 & 0 \end{bmatrix} \preceq \begin{bmatrix} X_{11}^1 & X_{12}^1 \\ X_{21}^1 & X_{22}^1 \end{bmatrix} \\ \begin{bmatrix} Y^2 & 0 \\ 0 & 0 \end{bmatrix} \preceq \begin{bmatrix} X_{11}^2 & X_{12}^2 \\ X_{21}^2 & X_{22}^2 \end{bmatrix} \end{matrix}$$

$$\rightsquigarrow (X,Y) = \lambda (X^1,Y^1) + (1-\lambda)(X^2,Y^2)$$

$$\rightarrow \lambda \begin{bmatrix} Y^1 & 0 \\ 0 & 0 \end{bmatrix} + (1-\lambda) \begin{bmatrix} Y^2 & 0 \\ 0 & 0 \end{bmatrix} \preceq \lambda \begin{bmatrix} X_{11}^1 & X_{12}^1 \\ X_{21}^1 & X_{22}^1 \end{bmatrix} + (1-\lambda) \begin{bmatrix} X_{11}^2 & X_{12}^2 \\ X_{21}^2 & X_{22}^2 \end{bmatrix}$$

$$\longrightarrow \begin{bmatrix} Y & 0 \\ 0 & 0 \end{bmatrix} \preceq \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}$$

بدیهتاً $Y \succ 0$, $X \succ 0$ $\leftarrow$ به ازای (X^1,Y^1) و (X^2,Y^2) ، (X,Y) هم عضو dom G هست، هر ترکیب convex ای از این دو هم در dom G است یعنی dom G مجموعه ای convex است.

پس G تابعی convex است، حالا طبق اسلاید های قبل اگر روی یکی از المان هایش inf بگیریم همچنان convex باشد $\leftarrow$ از طرفی داریم که:

$$f(X) = \inf_{Y} G(X,Y)$$

$X_{11} - X_{12}X_{22}^{-1}X_{21}$ جواب مسئله ۶ بهینه سازی بخش ب بود که از کمینه کردن G بدست می آید

چون G روی X, Y convex است

$\leftarrow$ $f(X)$ هم تابعی convex است.

ب)

از راهنمایی سوال استفاده می‌کنیم:

$$\begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}^{-1} = \begin{bmatrix} 0 & 0 \\ 0 & X_{22}^{-1} \end{bmatrix} + \begin{bmatrix} -I \\ X_{22}^{-1}X_{21} \end{bmatrix} (X_{11} - X_{12}X_{22}^{-1}X_{21})^{-1} \begin{bmatrix} -I \\ X_{22}^{-1}X_{21} \end{bmatrix}$$

حالا از دو طرف P^T و P را ضرب می‌کنیم، داریم که:

$$P^T X^{-1} P = P^T \left(\begin{bmatrix} 0 & 0 \\ 0 & X_{22}^{-1} \end{bmatrix} + \begin{bmatrix} -I \\ X_{22}^{-1}X_{21} \end{bmatrix} (X_{11} - X_{12}X_{22}^{-1}X_{21})^{-1} \begin{bmatrix} -I \\ X_{22}^{-1}X_{21} \end{bmatrix}^T \right) P =$$

$$[I \ \ 0] \left(\begin{bmatrix} 0 & 0 \\ 0 & X_{22}^{-1} \end{bmatrix} + \begin{bmatrix} -I \\ X_{22}^{-1}X_{21} \end{bmatrix} (X_{11} - X_{12}X_{22}^{-1}X_{21})^{-1} \begin{bmatrix} -I \\ X_{22}^{-1}X_{21} \end{bmatrix} \right) \begin{bmatrix} I \\ 0 \end{bmatrix} =$$

$$\underbrace{[I \ \ 0] \begin{bmatrix} -I \\ X_{22}^{-1}X_{21} \end{bmatrix}}_{-I} (X_{11} - X_{12}X_{22}^{-1}X_{21})^{-1} \underbrace{\begin{bmatrix} -I \\ X_{22}^{-1}X_{21} \end{bmatrix}^T \begin{bmatrix} I \\ 0 \end{bmatrix}}_{-I} =$$

$$(X_{11} - X_{12}X_{22}^{-1}X_{21})^{-1} \Rightarrow \boxed{P^T X^{-1} P = (X_{11} - X_{12}X_{22}^{-1}X_{21})^{-1}}$$

← حکم برقرار است.

a-7) داریم که:

$$d^{mp}(p,q) = \max\{|\underbrace{Prob(S;p) - Prob(S;P)}_{\sum_{i\in S} p_i - q_i}| \mid S \subseteq \{1,\dots,n\}\}$$

$$\rightarrow d^{mp}(p,q) = \max\{|\sum_{i\in S} p_i - q_i| \mid S \subseteq \{1,\dots,n\}\}$$

منطقاً ماکزیمم تابع زمانی اتفاق می‌افتد که S به گونه‌ای باشد که تمام $p_i - q_i$ های مثبت را فقط شامل شود یا فقط $p_i - q_i$ های منفی را شامل شود، زیرا در غیر اینصورت قدر مطلق مقدارش کاهش می‌یابد.

تمام اندیس‌هایی که $p_i - q_i \geq 0$ است را S_+ و آنهایی که $p_i - q_i < 0$ است را S_- می‌گیریم.

حالا داریم که:

$$\sum_{i\in S_+} p_i - q_i = \sum_{i\in S_+} p_i - \sum_{i\in S_+} q_i = \left(1 - \sum_{i\in S_-} p_i\right) - \left(1 - \sum_{i\in S_-} q_i\right)$$

$$= -\left(\sum_{i\in S_-} p_i - q_i\right) \longrightarrow \left|\sum_{i\in S_+} p_i - q_i\right| = \left|\sum_{i\in S_-} p_i - q_i\right|$$

داریم: $$\|p-q\|_1 = \sum_{i\in\{1,\dots,n\}} |p_i - q_i| = \underbrace{\left|\sum_{i\in S_+} p_i - q_i\right| + \left|\sum_{i\in S_-} p_i - q_i\right|}_{\text{برابر اند}}$$

$$\longrightarrow \underbrace{\underbrace{\left|\sum_{i\in S_+} p_i - q_i\right|}_{} = \underbrace{\left|\sum_{i\in S_-} p_i - q_i\right|}_{}}_{d^{mp}(p,q)} = \frac{1}{2}\|p-q\|_1$$

$$\rightarrow d^{mp}(p,q) = \frac{1}{2}\|p-q\|_1$$

نرم هم که تابعی convex است

$\leftarrow d^{mp}$ تابعی محدب است

Scanned with CamScanner

(b

$$d^{he}(p,q) = \sum_{i=1}^{n}(\sqrt{p_i}-\sqrt{q_i})^2 = \sum_{i=1}^{n} p_i+q_i-2\sqrt{p_iq_i} = 2-2\sum_{i=1}^{n}\sqrt{p_iq_i}$$

حالا باید نشان دهیم که $\sum_{i=1}^{n}\sqrt{p_iq_i}$ تابعی concave است یعنی عبارت زیر باید برقرار باشد،

$$\forall\lambda\in[0,1],\ (p^{(1)},q^{(1)}),\ (p^{(2)},q^{(2)}) \ :\ (p,q)=\lambda(p^{(1)},q^{(1)})+(1-\lambda)(p^{(2)},q^{(2)})$$

$$\longrightarrow \lambda\sum_{i=1}^{n}\sqrt{p_i^{(1)}q_i^{(1)}}+(1-\lambda)\sum_{i=1}^{n}\sqrt{p_i^{(2)}q_i^{(2)}} \le \sum_{i=1}^{n}\sqrt{(\lambda p_i^{(1)}+(1-\lambda)p_i^{(2)})(\lambda q_i^{(1)}+(1-\lambda)q_i^{(2)})}$$

این عبارت همواره باید برقرار باشد. ابتدا نشان می دهیم که به ازای هر (p_1,q_1) و (p_2,q_2) و $\lambda\in[0,1]$ داریم که،

$$\star\star\quad \sqrt{(\lambda p_1+(1-\lambda)p_2)(\lambda q_1+(1-\lambda)q_2)} \ge \lambda\sqrt{p_1q_1}+(1-\lambda)\sqrt{p_2q_2} \leftrightarrow$$

$$(\lambda p_1+(1-\lambda)p_2)(\lambda q_1+(1-\lambda)q_2) \ge \lambda^2p_1q_1+(1-\lambda)^2p_2q_2+2\lambda(1-\lambda)\sqrt{p_1q_1p_2q_2} \leftrightarrow$$

$$\cancel{\lambda^2p_1q_1}+\cancel{(1-\lambda)^2p_2q_2}+\lambda(1-\lambda)(p_1q_2+p_2q_1) \ge \cancel{\lambda^2p_1q_1}+\cancel{(1-\lambda)^2p_2q_2}+2\lambda(1-\lambda)\sqrt{p_1q_1p_2q_2}$$

$$\leftrightarrow p_1q_2+p_2q_1 \ge 2\sqrt{p_1q_2p_2q_1} \leftrightarrow (p_1q_2+p_2q_1)^2 \ge 4p_1q_2p_2q_1 \leftrightarrow$$

$$(p_1q_2-p_2q_1)^2 \ge 0 \longrightarrow$$

عبارت بالا برقرار است. $\star\star$

حالا داریم که،

$$\forall i:\quad \lambda\sqrt{p_i^{(1)}q_i^{(1)}}+(1-\lambda)\sqrt{p_i^{(2)}q_i^{(2)}} \le \sqrt{(\lambda p_i^{(1)}+(1-\lambda)p_i^{(2)})(\lambda q_i^{(1)}+(1-\lambda)q_i^{(2)})}$$

$$\longrightarrow \sum_{i=1}^{n}\lambda\sqrt{p_i^{(1)}q_i^{(1)}}+\sum_{i=1}^{n}(1-\lambda)\sqrt{p_i^{(2)}q_i^{(2)}} \le \sum_{i=1}^{n}\sqrt{(\lambda p_i^{(1)}+(1-\lambda)p_i^{(2)})(\lambda q_i^{(1)}+(1-\lambda)q_i^{(2)})}$$

$$\rightarrow \lambda \sum_{i=1}^{n} \sqrt{p_i^{(1)} q_i^{(1)}} + (1-\lambda) \sum_{i=1}^{n} \sqrt{p_i^{(2)} q_i^{(2)}} \le \sum_{i=1}^{n} \sqrt{(\lambda p_i^{(1)} + (1-\lambda) p_i^{(2)})(\lambda p_i^{(2)} + (1-\lambda) q_i^{(2)})}$$

→ این تابع concave است ← قرینه آن convex است.

← $2-2\sum_{i=1}^{n}\sqrt{p_i q_i}$ ، تابعی convex است ← $d^{he}(p,q)$ is convex

ادامه: حالا به بخش بعد می‌پردازیم: می‌خواهیم ثابت کنیم که $d^{kl}(p,q)$ ، تابعی convex است.

$$d^{kl}(p,q) = \sum_{i=1}^{n} p_i \log (p_i/q_i)$$

اگر به عبارات داخل سیگما نگاه کنیم، هرکدام از اعمال perspective روی تابع Ln بدست آمده اند.

← چون که $\log q_i$ عبارتی concave است، $p_i \log q_i/p_i$ نیز عبارتی concave است زیرا perspective تغییری در convexity ، concavity یک تابع نمی‌دهد.

← $-p_i \log q_i/p_i$ تابعی convex است ← $p_i \log \frac{p_i}{q_i}$ تابعی convex است.

← جمع یک سری تابع هم تابعی convex است ← $d^{kl}(p,q)$ تابعی convex است.

* $f: R^n \longrightarrow R$ is symmetric if $f(x) = f(Px)$ for all permutation matrices

$P \rightsquigarrow P^T 1 = 1 \;,\; P1 = 1$

* $f: S^n \longrightarrow R$ is symmetric invariant if for all orthogonal $Q \in R^{n \times n}$, $Q^T Q = I$

we have $f(QXQ^T) = f(X)$

$\lambda(X) \leftarrow$ برداری شامل مقادیر ویژه های X.

حالا باید نشان دهیم که $F: S^n \longrightarrow R$ یک unitary invariant است اگر و تنها اگر

یک f symmetric ای داشته باشیم که $f: R^n \longrightarrow R$ که $F(X) = f(\lambda(X))$

ابتدا طرف دوم را نشان می دهیم: اگر یک f symmetric ای داشته باشیم $F(X) = f(\lambda(X))$، آنگاه

F یک unitary invariant است، for all orthogonal Q, $Q^T Q = I$:

$$F(X) = f(\lambda(X)) \Rightarrow F(QXQ^T) = f(\lambda(QXQ^T))$$

$$Q \Lambda Q^T - \lambda I = Q(\Lambda - \lambda I) Q^T \rightsquigarrow$$ مقادیر ویژه X, QXQ^T یکسان است.

$$\Rightarrow \underbrace{f(\lambda(QXQ^T))}_{F(QXQ^T)} = \underbrace{f(\lambda(X))}_{F(X)} \longrightarrow \boxed{F(X) = F(QXQ^T)}$$

حالا به اثبات طرف دیگر می پردازیم:

اگر $F(X) = F(QXQ^T)$ (F یک unitary invariant باشد) آنگاه یک symmetric وجود دارد که:

$F(X) = f(\lambda(X))$.

از آنجایی که $X \in S^n \leftarrow X = Q \Lambda Q^T$.

$$F(X) = F(Q\Lambda Q^T) = F(Q^T Q \Lambda Q^T Q) = F(\Lambda(X)) \longrightarrow$$

$\Lambda(X)$ ماتریسی است که قطری است و روی قطر آن مقادیر ویژه X چیده شده اند $\longleftarrow$

$F(\Lambda(X))$ چون فقط تابع مقادیر ویژه X است، آنرا می توانیم به صورت $f(\lambda(X))$ نوشت که

g تابعی از $R^n \to R$ است.

حالا باید نشان دهیم که g تابعی symmetric است:

$F(X) = g(\lambda(X))$

این ضرب ماتریسی unitary است

$$X = Q\Lambda Q^T = Q \underbrace{P^T P}_{I} \Lambda \underbrace{P^T P}_{I} Q^T = (PQ^T)^T P\Lambda P^T (PQ^T)$$

$\leftarrow (PQ^T)^T PQ^T = QP^T PQ^T = QQ^T = I$

$$\rightsquigarrow F(X) = F(\Lambda) = F(P\Lambda P^T) = g(\lambda(X))$$

ماتریس قطری با درایه‌های $v_1 - v_n$ روی قطر اصلی

به ازای هر بردار v که به g ورودی داده شود، داریم:

$$\forall v \in \text{dom}\, g, \quad g(v) = F(\text{diag}(v)) = F(\underbrace{P\,\text{diag}(v)\,P^T})$$

همان $\text{diag}(v)$ است که درایه‌هاش روی قطر اصلی جابه‌جا شده اند.

در واقع عبارت است از $\text{diag}(Pv)$

$$\rightsquigarrow g(v) = F(\text{diag}(v)) = F(P\,\text{diag}(v)\,P^T) = F(\text{diag}(Pv)) = g(\lambda(\text{diag}(Pv)))$$

$$= g(Pv) \longrightarrow \forall v \in \text{dom}\, g, \quad g(v) = g(Pv)$$

← اگر $F(X)$ یک unitary invariant باشد، آنگاه یک symmetric f داریم که $F(X) = f(\lambda(X))$.

در نتیجه دو طرف رابطه را نشان دادیم و حکم مسئله برقرار است

(b

$f_{sy}(Y) = f(\lambda(Y))$

$f^*_{sy}(Y)$ عبارت است از اینکه از $f_{sy}(Y)$ ، conjugate بگیریم

$(f^*)_{sy}(Y)$ عبارت است از اینکه ابتدا از f conjugate بگیریم، سپس sy را اعمال کنیم.

$$\rightsquigarrow f_{sy}(Y): \quad f^*_{sy}(Y) = \sup_{X\in S^n} \{tr(XY) - f_{sy}(X)\}$$

$$\underbrace{(f^*)}_{g}{}_{sy}(Y) = g_{sy}(Y) = g(\lambda(Y)) = (f^*)(\lambda(Y)) = \sup_{v} \{v^T\lambda(Y) - f(v)\}$$

حالا باید نشان دهیم این دو عبارت معادل همدیگر هستند.

طبق راهنمایی سوال که در تمرین ۱ نیز اثبات شد، داریم که:

$$\forall X: \ tr(XY) \le \overbrace{\sum_{i=1}^{n} \lambda_i(X)\lambda_i(Y)}^{\lambda(X)^T\lambda(Y)}$$

حالا اگر از دو طرف ماکزیمم بگیریم روی X، باز هم نامساوی برقرار است ←

$$\sup_{X\in S^n} tr(XY) \le \boxed{\sup_{X} \lambda(X)^T\lambda(Y)} = \sup_{\lambda(X)} \lambda(X)^T\lambda(Y) = \sup_{v} v^T\lambda(Y)$$

از آنجایی که دامنه ماتریس ها در این سوال، ماتریس های متقارن است، برای ماتریس های متقارن رابطه را بازنویسی می کنیم.

$Y\in S^n \longrightarrow Y = Q\Lambda(Y)Q^T \longrightarrow$ اگر X را اینگونه انتخاب کنیم $Q\Lambda(X)Q^T$ ، آنگاه داریم که:

$$tr(XY) = tr(Q\Lambda(Y)\overset{I}{\cancel{Q^TQ}}\Lambda(X)Q^T) = tr(\overset{I}{\cancel{Q^TQ}}\Lambda(Y)\Lambda(X)) = tr(\Lambda(X)\Lambda(Y)) =$$

$\lambda(X)^T\lambda(Y) \longrightarrow$ در این حالت $tr(XY)$ با کران بالایی برابر می شود

if we choose $X = Q\Lambda(X)Q^T \rightsquigarrow * \ tr(XY) = \lambda^T(X)\lambda(Y)$

$\rightarrow$ if we choose $X = Q\Lambda(X)Q^T \longrightarrow \sup_{\Lambda(X)} tr(XY) = \sup_{\lambda(X)} \lambda(X)^T\lambda(Y)$

$$\sup_{X\in S^n}\{tr(XY)\} = \sup_{X\in S^n}\{\lambda(X)^T\lambda(Y)\} = \sup_{v}\{v^T\lambda(Y)\} \longleftarrow$$

$$\rightsquigarrow \; f^{\star}_{sy}(Y) = \sup_{X}\{tr(XY) - f_{sy}(X)\} = \sup_{X}\{tr(XY) - f(\lambda(X))\}$$

$$= \sup_{X}\{\lambda(X)^T\lambda(Y) - f(\lambda(X))\} = \sup_{\lambda(X)}\{\lambda(X)^T\lambda(Y) - f(\lambda(X))\} =$$

$$\sup_{v}\{v^T\lambda(Y) - f(v)\} = (f^{\star})_{sy}(Y) \Rightarrow$$

$$\underline{f^{\star}_{sy}(Y) = \sup_{X}\{tr(XY) - f_{sy}(Y)\} = (f^{\star})_{sy}(Y)}$$

(c

f is symmetric, closed convex $\xrightarrow{?}$ $f_{sy}(X) = f(\lambda(Y))$ is also closed convex, unitary invariant

$f(x) = f(Px)$

$$f_{sy}(X) = f(\lambda(X)) \rightarrow f_{sy}(QXQ^T) = f(\lambda(QXQ^T)) = f(\lambda(X)) = f_{sy}(X)$$

$\longrightarrow f_{sy}(QXQ^T) = f_{sy}(X) \longrightarrow$ <u>f_{sy} is unitary invariant.</u>

حالا می دانیم که conjugate هر تابع convex , closed است.

همچنین داشتیم که $f^*_{sy}(Y) = (f^*)_{sy}(Y)$ به ازای هر تابع f که convex , symmetric است، برقرار است.

حالا نشان می دهیم که f^* هم symmetric است و از طرفی می دانیم که convex است ← خواهیم داشت که $(f^{**})_{sy}(Y) = (f^*)^*_{sy}(Y)$ ، از طرفی هم چون $f^*_{sy}(Y) = (f^*)_{sy}(Y) \leftarrow f^{**}_{sy}(Y) = (f^{**})^*_{sy}(Y)$.

زیرا داریم که:

$$f^{**}_{sy}(Y) = \sup_X \{tr(XY) - f^*_{sy}(X)\} = \sup_X \{tr(XY) - (f^*)_{sy}(X)\} = (f^*)^*_{sy}(Y)$$

حالا باید نشان دهیم که f^* هم symmetric است.

$f(v) = f(Pv) \leftarrow$ symmetric f است ←

$$(f^*)_{sy}(Y) = f^*(\lambda(Y)) = \sup_v \{ v^T \underset{I = P^T P}{\wedge} \lambda(Y) - f(v)\} = \sup_v \{v^T P^T P \lambda(Y) - f(Pv)\}$$

چون به ازای هر P ای برقرار است ← همچنین تمام v ها در sup هستند.

$$= \sup_v \{(Pv)^T P \lambda(Y) - f(\underbrace{Pv}_{u})\} = \sup_u \{u^T (P\lambda(Y)) - f(u)\} = f^*(P\lambda(Y))$$

$\rightarrow f^*(\lambda(Y)) = f^*(P\lambda(Y))$ ، همچنین اثبات قبل کلی بود و به جای $\lambda(Y)$ هر چیزی قرار دهیم باز هم $f^*(v) = f^*(Pv)$.

$$f^{**}(\lambda(Y)) = (f^{**})_{sy}(Y) = (f^*)^*_{sy}(Y)$$ ← پس f^* هم یک symmetric است

$$= f^{**}_{sy}(Y)$$

پس از آنجایی که f ، convex , closed بود ← ★ f^{**} همان f است ←

$$f^{**}_{sy}(Y) = f_{sy}(Y) \Rightarrow$$

از آنجایی که conjugate یک تابع convex , closed است ←

conjugate conjugate یک تابع هم convex است ← $f^{**}_{sy}(Y)$ تابعی convex است

← $f_{sy}(Y)$ نیز تابعی convex است.

← اگر f تابعی convex و closed باشد ← $f_{sy}(Y)$ نیز تابعی convex, closed, unitary invariant است.

d) F تابعی convex و closed است ← به‌ازای دو ماتریس متقارن Λ_1 و Λ_2 با مقادیر ویژه به صورت بردار λ_1 و λ_2، داریم که،

$$F(\Lambda_1) = f(\lambda_1)$$
$$F(\Lambda_2) = f(\lambda_2) \rightsquigarrow \underbrace{F(\theta\Lambda_1 + (1-\theta)\Lambda_2)}_{\underbrace{f(\lambda(\theta\Lambda_1 + (1-\theta)\Lambda_2))}_{f(\theta\lambda_1 + (1-\theta)\lambda_2)}} \le \theta \underbrace{F(\Lambda_1) + (1-\theta) F(\Lambda_2)}_{\theta f(\lambda_1) + (1-\theta) f(\lambda_2)}$$

$$\Rightarrow f(\theta\lambda_1 + (1-\theta)\lambda_2) \le \theta f(\lambda_1) + (1-\theta) f(\lambda_2) \rightarrow$$ f هم تابعی convex است.

همچنین f symmetric هم بود ← f symmetric را داریم که convex است.

این مسئله نسبت به مسئله ی بخش الف تنها convex بودن f را برای شرط و اثبات داراضافه تر دارد در نتیجه تنها باید ثابت کنیم که این تابع convex است

e) if f is symmetric and convex:

$$\text{tr}(GX) = \text{tr}(U\,\text{diag}(g)\,\underbrace{U^T U}_{I}\,\text{diag}(x)\,U^T) \rightarrow \text{tr}(\text{diag}(g)\,\text{diag}(x))$$

$$f_{sy}(X) + \text{tr}(G(Y-X)) = f(\lambda(x)) + \text{tr}(GY) - \underbrace{\text{tr}(GX)}_{g^T x}$$

$$= f(\lambda(x)) + \underbrace{\text{tr}(GY)}_{\underbrace{\le \sum_{i=1}^{n} \lambda_i(G)\lambda_i(Y)}_{g^T\lambda(Y)}} - g^T x \le \underbrace{f(\lambda(x)) + g^T\lambda(Y) - g^T x}_{f(\lambda(x)) + g^T(\lambda(Y) - x) \le f(\lambda(Y))}$$

طبق سوال به ازای هر بردار در R^n برقرار است.

$$= f_{sy}(Y)$$

$$\rightarrow \underline{f_{sy}(X) + \text{tr}(G(Y-X)) \le f_{sy}(Y)} \rightarrow \text{if } f \text{ is symmetric and convex}$$

$$\Rightarrow \boxed{f_{sy}(Y) \ge f_{sy}(x) + \text{tr}(G(Y-X))}$$