به نام خدا

علی قاسم زاده

۴۰۱۱۰۶۳۳۹

بهینه سازی محدب

تمرین ۴

سوال اول

الف) داریم که :

$$Minimize \quad \|Ax-b\|_1$$

حالا بردار $Ax-b$ را در نظر بگیرید برای قدر مطلق تک تک درایه های آن یک کران بالای s_i در نظر می گیریم آنگاه داریم که

$$\forall_i : \quad |(Ax-b)_i| \leq s_i$$

پس داریم که :

$$\forall_i : \quad s_i \leq (Ax-b)_i \leq s_i$$

که این معادل این است که :

$$\forall_i (Ax-b)_i \leq s_i \,, \quad (-Ax+b)_i \leq s_i$$

حالا می توانیم s_i ها را در یک بردار قرار دهیم و آنرا بردار s بنامیم و قیود را به صورت برداری بنویسیم یعنی داریم که :

$$Ax-b \preceq s, \quad -Ax+b \preceq s$$

همچین $objective$ ما عبارت است از مینیمم کردن $\|s\|_1$ پس مسئله ی ما به صورت زیر نوشته می شود :

$$\begin{aligned} \min_{s,x} \quad & \|s\|_1 \\ \text{s.t.} \quad & Ax-b \preceq s \\ & -Ax+b \preceq s \end{aligned}$$

حال می توانیم آنرا به صورت زیر باز نویسی کنیم که یک مسئله ی LP است :

$$\begin{aligned} \min_{s,x} \quad & \mathbf{1}^T s \\ \text{s.t.} \quad & Ax-b \preceq s \\ & -Ax+b \preceq s \end{aligned}$$

چون هم قیود و هم تابع $objective$ ما خطی هستند پس مسئله ی ما یک مسئله ی LP است.

ب) داریم که :

$$Minimize \quad \|x\|_\infty$$

حالا یک کران بالا برای نرم بینهایت x در نظر می گیریم و آنرا مینیمایز می کنیم و x کوچکتر مساوی آن کران بالا را به قیود اضافه می کنیم داریم :

$$\begin{aligned} \min_{x,t} \quad & t \\ \text{s.t.} \quad & \|x\|_\infty \leq t \end{aligned}$$

از طرفی داریم که وقتی نرم بینهایت یک بردار کوچکتر مساوی مقداری باشند یعنی ماکزیمم بین عناصرش از آن مقدار کوچکتر مساوی است پس داریم که بردار x از بردار تماما t کوچکتر مساوی است :

$$\begin{aligned} \min_{x,t} \quad & t \\ \text{s.t.} \quad & x \preceq t\mathbf{1} \end{aligned}$$

حالا می خواهیم که تابع زیر را مینیمایز کنیم :

$$\|Ax - b\|_1 + \|x\|_\infty$$

از ترکیب عبارت بخش ب و عبارت بالا داریم که:

$$\begin{aligned} \min_{t,s,x} \quad & \mathbf{1}^T s + t \\ \text{s.t.} \quad & x \preceq t\mathbf{1} \\ & Ax - b \preceq s \\ & -Ax + b \preceq s \end{aligned}$$

عبارت بالا هم معادل یک مسئله ی LP است.

سوال دوم

داریم که :

$$max \quad min_{i=1,\dots,n} \quad \frac{S_i}{I_i + \sigma_i}$$

عبارت $t \quad min$ معادل این است که ما روی مقادیر $\frac{1}{t}$ یک max بگیریم و سپس حاصل را وارونه کنیم به عبارتی داریم که $min\ t$ معادل است با $\frac{1}{max\ \frac{1}{t}}$ حالا همین

را روی عبارت صورت سوال اعمال می کنیم :

$$max \quad min_{i=۱,...,n} \quad \frac{S_i}{I_i+\sigma_i} \leftrightarrow max \frac{۱}{max_{i=۱,...,n} \quad \frac{I_i+\sigma_i}{S_i}}$$

حالا باز هم برای max داریم که : $max \frac{۱}{t} = \frac{۱}{min\ t}$ پس داریم که : عبارت صورت سوال معادل این است که :

$$max \quad min_{i=۱,...,n} \quad \frac{S_i}{I_i+\sigma_i} \leftrightarrow \frac{۱}{min \quad max_{i=۱,...,n} \frac{I_i+\sigma_i}{S_i}}$$

این هم یعنی اینکه ما داریم

$$min \quad max_{i=۱,...,n} \frac{I_i+\sigma_i}{S_i}$$

می گیریم و سپس معکوس آن را به عنوان $objective$ ارائه می دهیم پس عبارت صورتت سوال معادل $min \quad max_{i=۱,...,n} \frac{I_i+\sigma_i}{S_i}$ است حالا داریم که :

$$max \quad min_{i=۱,...,n} \quad \frac{S_i}{I_i+\sigma_i} \leftrightarrow min \quad max_{i=۱,...,n} \frac{I_i+\sigma_i}{S_i}$$

پس مسئله ی ما با توجه به قید هایی که در صورت سوال ذکر شده اند عبارت است از :

$$\begin{aligned} \min_x \quad & max_{i=۱,...,n} \frac{\sum_{k\neq i} G_{ik}p_k+\sigma_i}{G_{iip_i}} \\ \text{s.t.} \quad & ۰ \leq p_i \leq P_i^{max} \quad i=۱,...,n \\ & \sum_{k\in K_l} p_k \leq P_l^{gp} \quad i=۱,...,m \\ & \sum_{k=۱}^{n} G_{ik}p_k \leq P_i^{rc} \quad i=۱,...,n \end{aligned}$$

که این هم فرم $generalizedlinear - fractional$ است.

سوال سوم

الف) داریم که :

$$\begin{aligned} \min_{x} \quad & \sup_{P\in\mathcal{E}} (\frac{۱}{۲}x^TPx + q^Tx + r) \\ \text{s.t.} \quad & Ax \preceq b \end{aligned}$$

همچنین داریم که مجموعه ی $\mathcal{E}$ عبارت است از $\{P_۱, ..., P_k\}$
یک کران بالا برای $\sup_{P\in\mathcal{E}} (\frac{۱}{۲}x^TPx + q^Tx + r)$ می گذاریم و آنرا t می نامیم، خواهیم داشت که :

$$\begin{aligned} \min_{t,x} \quad & t \\ \text{s.t.} \quad & Ax \preceq b \\ & \frac{۱}{۲}x^TPx + q^Tx + r \leq t, \quad i = ۱...k \\ & \|y\|_۲ \leq s \end{aligned}$$

که این فرم هم یک مسئله ی $QCQP$ است.

ب) داریم که :

$$\gamma I \preceq P - P_۰ \preceq \gamma I \quad \longrightarrow \quad P_۰ - \gamma I \preceq P \preceq P_۰ + \gamma I \longrightarrow$$
$$P - P_۰ - \gamma I \preceq ۰ \longrightarrow x^T(P - P_۰ - \gamma I)x \leq ۰ \longrightarrow$$
$$x^TPx \leq x^TP_۰x + \gamma x^TIx \longrightarrow x^TPx \leq x^TP_۰x + \gamma x^Tx \longrightarrow$$

مینیمایز کردن sup عبارت معادل مینیمایز کردن کران بالای عبارت است زیرا قبول وصول است. پس خواهیم داشت :

$$\begin{aligned} \min_{x} \quad & x^T(P_۰ + \gamma I)x \\ \text{s.t.} \quad & Ax \preceq b \end{aligned}$$

این مسئله هم نسبت به x یک مسئله ی QP است.

پ) داریم که :

$$P = P_۰ + \sum_{i=۱}^{k} P_i u_i$$

و مسئله ی اصلی عبارت است از :

$$\begin{array}{ll} \min\limits_{x} & \sup\limits_{P\in\mathcal{E}} (\frac{۱}{۲}x^TPx+q^Tx+r) \\ \text{s.t.} & Ax\preceq b \end{array}$$

حال آنرا به صورت زیر می نویسیم، داریم که :

$$\begin{array}{ll} \min\limits_{x} & \sup\limits_{P\in\mathcal{E}} (\frac{۱}{۲}x^T(P_۰+\sum_{i=۱}^{k}P_iu_i)x+q^Tx+r) \\ \text{s.t.} & Ax\preceq b \end{array}$$

حالا عبارت بالا را بسط می دهیم :

$$\begin{array}{ll} \min\limits_{x} & \sup\limits_{P\in\mathcal{E}} (\frac{۱}{۲}x^TP_۰x+\frac{۱}{۲}x^T(\sum_{i=۱}^{k}P_iu_i)x+q^Tx+r) \\ \text{s.t.} & Ax\preceq b \end{array}$$

داریم که :

$$\begin{array}{ll} \min\limits_{x} & \sup\limits_{P\in\mathcal{E}} (\frac{۱}{۲}x^TP_۰x+\frac{۱}{۲}\sum_{i=۱}^{k}u_ix^TP_ix+q^Tx+r) \\ \text{s.t.} & Ax\preceq b \end{array}$$

از طرفی داریم که :

$$\sum_{i=۱}^{k}u_ix^TP_ix=[u_۱,...,u_k]\begin{bmatrix}x^TP_۱x\\ \vdots \\ x^TP_kx\end{bmatrix}\longrightarrow$$

بردار دومی را بردار v می نامی داریم :

$$\sum_{i=۱}^{k}u_ix^TP_ix=u^Tv\leq\|u\|_۲\,\|v\|_۲\leq\|v\|_۲=\left\|\begin{bmatrix}x^TP_۱x\\ \vdots \\ x^TP_kx\end{bmatrix}\right\|_۲$$

حالا یک کران بالا برای تمام $\frac{1}{2}x^TP_ix$ ها میزاریم به اسم y_i و آنرا مینیمایز می کنیم،
چون P_i ها عضو S_+^n هستند پس داریم :

$$\forall_i \quad : \quad x^TP_ix \geq 0$$

پس تمامی y_i ها نیز مثبت اند حالا داریم که : $\frac{1}{2}\|q\|_2 \leq \|y\|_2$ پس خواهیم داشت

$$\frac{1}{2}\sum_{i}^{k} u_i x^TP_ix \leq \frac{1}{2}\|v\|_2 \leq \|y\|_2$$

حال مسئله را به صورت زیر باز نویسی می کنیم :

$$\begin{aligned} \min_{x,y} \quad & (\frac{1}{2}x^TP_0x + \|y\|_2 + q^Tx + r) \\ \text{s.t.} \quad & Ax \preceq b \\ & \frac{1}{2}x^TP_ix \leq y_i, \quad i = 1, ..., k \end{aligned}$$

حالا یک کران بالا روی $\frac{1}{2}x^TP_0x$ به اسم t در نظر می گیریم خواهیم داشت که :

$$\frac{1}{2}x^TP_0x \leq t$$

حالا داریم که :

$$\frac{1}{2}x^TP_0x \leq t \longrightarrow \frac{1}{2}x^TP_0^{\frac{1}{2}}P_0^{\frac{1}{2}}x \leq t \longrightarrow x^TP_0^{\frac{1}{2}}P_0^{\frac{1}{2}}x \leq 2t \longrightarrow$$

$$x^TP_0^{\frac{1}{2}}P_0^{\frac{1}{2}}x + (t - \frac{1}{2})^2 \leq (t + \frac{1}{2})^2 \longrightarrow \left\| \begin{bmatrix} P_0^{\frac{1}{2}}x \\ t - \frac{1}{2} \end{bmatrix} \right\|_2^2 \leq (t + \frac{1}{2})^2$$

$$\frac{1}{2}x^TP_ix \leq y_i \longrightarrow \frac{1}{2}x^TP_i^{\frac{1}{2}}P_i^{\frac{1}{2}}x \leq y_i \longrightarrow x^TP_i^{\frac{1}{2}}P_i^{\frac{1}{2}}x \leq 2y_i \longrightarrow$$

$$x^TP_i^{\frac{1}{2}}P_i^{\frac{1}{2}}x + (y_i - \frac{1}{2})^2 \leq (y_i + \frac{1}{2})^2 \longrightarrow \left\| \begin{bmatrix} P_i^{\frac{1}{2}}x \\ y_i - \frac{1}{2} \end{bmatrix} \right\|_2^2 \leq (y_i + \frac{1}{2})^2$$

همچنین برای $\|y\|_۲$ نیز یک کران بالا می گذاریم به اسم s حالا خواهیم داشت :

$$\begin{aligned} \min_{u,t,x} \quad & t+s+r \\ \text{s.t.} \quad & Ax \preceq b \\ & \left\| \begin{bmatrix} P^{\frac{۱}{۲}}x \\ t-\frac{۱}{۲} \end{bmatrix} \right\|_۲ \leq (t+\frac{۱}{۲}) \\ & \left\| \begin{bmatrix} P_i^{\frac{۱}{۲}}x \\ y_i-\frac{۱}{۲} \end{bmatrix} \right\|_۲ \leq (y_i+\frac{۱}{۲}), \quad i=۱,...,k \end{aligned}$$

این عبارت هم فرم $SOCP$ مسئله ی بالا است .

سوال چهارم

الف) به ازای هر بردار u داریم که :

$$\lambda_{max}(A,B) \leq t \longrightarrow \sup_u \frac{u^TAu}{u^TBu} \leq t$$

از طرفی داریم که $B \in S^n_{++}$ پس داریم که به ازای هر u : $u^TBu > ۰$ حالا طرفین را در $u^TBu > ۰$ ضرب می کنیم، داریم که :

$$u^TAu \leq tu^TBu \leq ۰ \longrightarrow u^T(A-tB)u \leq ۰$$

این رابطه همواره برقرار است پس ماتریس $A-tB$ یک ماتریس نیمه معین منفی است (بخش ب)
این هم معادل این است که تمامی مقادیر ویژه ی ماتریس $A-tB$ منفی باشند یا به عبارتی مقدار ویژه ی ماکزیمم نا مثبت باشد پس خواهیم داشت که عبارت صورت سوال معادل است با :

$$\lambda_{max}(A-tB) \leq ۰$$

پس تابع $\phi_t(A,B)$ را همان $\lambda_{max}(A-tB)$ در نظر می گیریم.

ب) در بخش قبل داشتیم که به ازای هر u ، $u^T(A-tB)u \leq ۰$ برقرار است یعنی داریم $A-tB \preceq ۰$
پس تابع $\phi_t(A,B)$ را همان ماتریس $A-tB$ در نظر می گیریم.

پ) داریم که :

$$\lambda_{max}(A(x),B(X)) = \sup_u \frac{u^TA(x)u}{u^TB(x)u} = \sup_u \frac{u^TA(x)u}{u^T(a^Tx+b)Iu} =$$

$$\sup_u \frac{u^TA(x)u}{(a^Tx+b)u^Tu} = \frac{\lambda_{max}(A(x))}{a^Tx+b}$$

حالا $\frac{۱}{a^Tx+b}$ را s می نامیم خواهیم داشت که :

$$\begin{aligned} \min_{x,s} \quad & s\lambda_{max}(A_۰ + x_۱A_۱ + x_۲A_۲ + ... + x_nA_n) \\ \text{s.t.} \quad & s > ۰ \end{aligned}$$

همچنین sx_i ها را y_i می نامیم، در نتیجه خواهیم داشت که :

$$y_i = sx_i \to y = sx \to a^Ty = sa^Tx \to a^Ty + bs = sa^Tx + bs \to$$

$$a^Ty + bs = (a^Tx+b)s \to a^Ty + bs = \frac{۱}{a^Tx+b}(a^Tx+b) = ۱$$

پس مسئله را به صورتت زیر بازنویسی می کنیم :

$$\begin{aligned} \min_{y,s} \quad & \lambda_{max}(sA_۰ + y_۱A_۱ + y_۲A_۲ + ... + y_nA_n) \\ \text{s.t.} \quad & s > ۰ \\ & a^Ty + bs = ۱ \end{aligned}$$

در نتیجه حکم مسئله اثبات می شود.

سوال پنجم

الف) برای $convex$ بودن عبارت اول مثال نقض ارائه می دهیم، x را تک متغیره در نظر می گیریم و A را $\begin{bmatrix} ۳ \\ ۰ \end{bmatrix}$ و b را $\begin{bmatrix} ۱ \\ ۰ \end{bmatrix}$ در نظر بگیریم، خواهیم داشت که :

$$\frac{\|Ax-b\|_۱}{۱-\|x\|_\infty} = \frac{|۳x+۱|}{۱-|x|}$$

اگر این تابع را در بازه ی $(-\frac{۱}{۲}, ۰)$ در نظر بگیریم خواهیم داشت که :

$$\frac{۳x+۱}{x+۱} = ۳ - \frac{۲}{۱+x} \xrightarrow{\frac{\partial}{\partial x}} \frac{۱}{(۱+x)^۲} \xrightarrow{\frac{\partial}{\partial x}} -\frac{۲}{(۱+x)^۳}$$

مشتق دوم در این بازه مثبت نیست پس تابع ما $convex$ نیست. حالا برای نشان دادن اینکه $quasiconvex$ است باید نشان دهیم که تمام $\alpha - sublevelset$ هایش مجموعه هایی $convex$ هستند.

$$S_\alpha = \{x \in R^n | f(x) \leq \alpha\}$$
$$\frac{\|Ax-b\|_۱}{۱ - \|x\|_\infty} = \frac{|۳x+۱|}{۱-|x|} \leq \alpha \rightarrow |۳x+۱| \leq \alpha(۱-|x|)$$

به ازای $\alpha \leq ۰$ که جوابی ندایم(مجموعه ی تهی)
حالا به ازای $\alpha \geq ۰$ داریم که

$$\|Ax-b\|_۱ + \alpha\|x\|_\infty \leq \alpha$$

که توابع $\|Ax-b\|_۱$ و $\alpha\|x\|_\infty$ هر دو توابعی محدب نسبت به x هستند در نتیجه جمعشان هم تابعی $convex$ است و یک $sublevelset$ از یک تابع $convex$ داریم که مجموعه ای $convex$ است، به ازای هر α این رابطه برقرار است پس تمام $sublevelset$ های ما مجموعه هایی $convex$ هستند. پس مسئله ی بخش اول یک مسئله ی $quasiconvex$ است.

حالا به برسی $convexity$ عبارت دوم می پردازیم :

$$\frac{\|Ax-b\|_۱^۲}{۱ - \|x\|_\infty} = \frac{|۳x+۱|}{۱-|x|}$$

می خواهیم اثبات کنیم که این تابع $convex$ است در نظر بگیرید تابع

$$\phi(s,t) = \frac{s^۲}{۱-t} \quad for\ s \leq ۰, \quad ۰\ for\ s\ <\ ۰$$

این تابع نسبت به المان های ورودی اش $convex$ و $nondecreasing$ همچنین توابع $\|x\|_\infty$ و $\|ax-b\|_۱^۲$ نیز توابعی $convex$ هستند پس با اعمال این تابع روی آن دو تابع، تابعی $convex$ حاصل می شود. در نتیجه تابع نهایی ما نیز $convex$ است.

ب) داریم که :

$$\min_x \quad \frac{\|Ax-b\|_۱}{۱ - \|x\|_\infty}$$

حالا یک کران بالا برای $\frac{1}{1-\|x\|_\infty}$ می گذاریم به اسم : $\frac{1}{t}$ حالا خواهیم داشت که :

$$\frac{1}{1-\|x\|_\infty} \leq \frac{1}{t} \to t \leq 1 - \|x\|_\infty \to \|x\|_\infty + t \leq 1$$

$$\begin{aligned} \min_{x,t} \quad & \frac{\|Ax-b\|_1}{t} \\ \text{s.t.} \quad & \|x\|_\infty + t \leq 1 \end{aligned}$$

حالا $\frac{x}{t}$ را y می نامیم و $\frac{1}{t}$ را نیزی s می نامیم پس با بردن t به داخل نرم و تقسیم قید بر t خواهیم داشت که :

$$\begin{aligned} \min_{y,s} \quad & \|Ay-bs\|_1 \\ \text{s.t.} \quad & \|y\|_\infty + 1 \leq s \end{aligned}$$

حالا مانند سوال اول تمرین یک کران بالا برای نرم ها قرار می دهیم و آنرا مینیمایز می کنیم و یک عبارت به قیود اضافه می کنیم:

$$min\ \|Ay-bs\|_1 \leftrightarrow min\ \mathbf{1}^T u\ s.t.\ -u \preceq Ay-bs \preceq u$$

$$\|y\|_\infty + 1 \leq s \leftrightarrow v+1 \leq s,\ -v\mathbf{1} \preceq y \preceq v\mathbf{1}$$

حالا مسئله را بازنویسی می کنیم و خواهیم داشت که :

$$\begin{aligned} \min_{u,v,y,s} \quad & \mathbf{1}^T u \\ \text{s.t.} \quad & -u\mathbf{1} \preceq Ay-bs \preceq u\mathbf{1} \\ & -v\mathbf{1} \preceq y \preceq v\mathbf{1} \\ & 1+v \leq s \end{aligned}$$

پس مسئله را می توانیم به فرم LP در آوریم و آنرا حل کنیم.

حالا به برسی راه حل تابع دومی می پردازیم :

$$\min_{x} \quad \frac{\|Ax-b\|_1^2}{1-\|x\|_\infty}$$

یک کران بالا روی $\|Ax-b\|_۱^۲$ و یک کران بالا روی $\|x\|_\infty$ می گذاریم،و عبارت حاصل را مینیمایز می کنیم با یه سری قیود اضافه خواهیم داشت که(دقت شود که t باید کمتر از ۱ باشد) :

$$\begin{aligned} \min_{x,s,t} \quad & \frac{s^۲}{۱-t} \\ \text{s.t.} \quad & \|Ax-b\|_۱ \leq s \\ & \|x\|_\infty \leq t \\ & t \leq ۱ \end{aligned}$$

حالا یک کران بالا برای $objective$ مان می گذاریم و آنرا مینیمایز می کنیم و یه سری قید به قیودمون اضافه می کنیم :) :

$$\begin{aligned} \min_{x,s,t,y} \quad & y \\ \text{s.t.} \quad & \|Ax-b\|_۱ \leq s \\ & \|x\|_\infty \leq t \\ & \frac{s^۲}{۱-t} \leq y \\ & t \leq ۱ \end{aligned}$$

حالا مانند سوال های قبل یک کران بالا برای تک تک عناصر $(Ax-b)_i$ در نظر می گیریم به اسم z_i و $\mathbf{۱}^T z$ را کمتر مساوی s قرار می دهیم همچنین داریم که :

$$\|x\|_\infty \leq t \leftrightarrow -t\mathbf{۱} \leq x \leq t\mathbf{۱}$$

پس مسئله ی ما به صورت زیر بازنویسی می شود :

$$\begin{aligned} \min_{x,s,t,y,z} \quad & y \\ \text{s.t.} \quad & -z \preceq Ax-b \preceq z \\ & \mathbf{۱}^T z \leq s \\ & -t\mathbf{۱} \preceq x \preceq t\mathbf{۱} \\ & t \leq ۱ \\ & \frac{s^۲}{۱-t} \leq y \end{aligned}$$

حالا قید آخر را به فرم نرم می بریم :

$$\frac{s^۲}{۱-t} \leq y \leftrightarrow s^۲ \leq y(۱-t) \leftrightarrow s^۲ \leq \frac{۱}{۴}\left(((۱-t)+y)^۲ - ((۱-t)-y)^۲\right) \leftrightarrow$$

$$s^۲ + \frac{۱}{۴}((۱-t)-y)^۲ \leq \frac{۱}{۴}((۱-t)+y)^۲ \leftrightarrow \left\|\begin{bmatrix} s \\ \frac{۱}{۲}(۱-t-y)\end{bmatrix}\right\|_۲^۲ \leq (\frac{۱}{۲}(۱-t+y))^۲$$

$$\leftrightarrow \left\|\begin{bmatrix} s \\ \frac{۱}{۲}(۱-t-y)\end{bmatrix}\right\|_۲ \leq \frac{۱}{۲}(۱-t+y)$$

حالا یکبار دیگر مسئله را بازنویسی می کنیم :

$$\begin{aligned} \min_{x,z,y,t,s} \quad & y \\ \text{s.t.} \quad & -z \preceq Ax-b \preceq z \\ & \mathbf{۱}^T z \leq s \\ & -t\mathbf{۱} \preceq x \preceq t\mathbf{۱} \\ & t \leq ۱ \\ & \left\|\begin{bmatrix} s \\ \frac{۱}{۲}(۱-t-y)\end{bmatrix}\right\|_۲ \leq \frac{۱}{۲}(۱-t+y) \end{aligned}$$

این مسئله هم یک مسئله ی $SOCP$ است که می توانیم آنرا با $solver$ حل کنیم.

سوال ششم

داریم که :

$$f(x) = (Ax+b)^T(P_۰ + x_۱P_۱ + ... + x_nP_n)^{-۱}(Ax+b) \quad P_i's\ are\ summetric$$

$$dom\, f = \{x \in R^n \,|\, P_۰ + x_۱P_۱ + ... + x_nP_n \succ ۰\}$$

حالا می خواهیم برای $epif$ یک LMI ارائه دهیم، می دانیم که :
از طرفی می دانیم $f(x) = (Ax+b)^T(P_۰ + x_۱P_۱ + ... + x_nP_n)^{-۱}(Ax+b) \geq ۰$
که t نیز مثبت است و باز هم داریم که : $(P_۰ + x_۱P_۱ + ... + x_nP_n) \succ ۰$ حالا اگر LMI متناظر با آن را می نویسیم :

$$\begin{bmatrix} t & (Ax+b)^T \\ (Ax+b) & P_۰ + x_۱P_۱ + ... + x_nP_n \end{bmatrix} \succeq ۰$$

نیمه مثبت معین بودن این ماتریس معادل است با اینکه (طبق تمرین اول که اثبات شد) :

$$t \geq 0 \quad \left|\begin{bmatrix} t & (Ax+b)^T \\ (Ax+b) & P_0 + x_1 P_1 + ... + x_n P_n \end{bmatrix}\right| \geq 0$$

$$(Ax+b)^T (P_0 + x_1 P_1 + ... + x_n P_n)^{-1} (Ax+b) \geq 0$$

این قیود هم همان قیود قبلی هستند پس این ماتریس معادل است با همان $F(x,t)$ همچنین $F(x,t)$ روی ورودی هایش $affine$ است.

سوال هفتم

الف) داریم که :

$$\begin{aligned} \min_x \quad & f_0(x) + \hat{g}(x; x^k) \\ \text{s.t.} \quad & f_i(x) \leq 0, \quad i = 1, ..., m \\ & Ax = b \end{aligned}$$

قیودمان که مشکلی ندارند تنها تابع $objective$ مشکل داشت که حالا باید بررسی کنیم و ببینیم که در حالت جدید تابع $objective$ تابعی $convex$ است یا خیر. تابع f که تابعی $convex$ است فقط باید تابع $\hat{g}$ را بررسی کنیم.

$$\hat{g}(x; x^k) = g(x^k) + \nabla g(x^k)^T (x - x^k) = g(x^k) - \nabla g(x^k)^T x^k + \nabla g(x^k)^T x$$

دو عبارت اولی که نسبت به x ثابت اند فقط می ماند عبارت سومی که آن هم نسبت به x تابعی $affine$ است حالا جمع این تابع با $f(x)$ که تابعی $convex$ است تابعی $convex$ خروجی می دهد. در نتیجه در هر $convex - concaveprocedure$ ای تابع $objecctive$ مان تابعی $convex$ است. قیود هم که مشکلی نداشتند پس در هر $convex - concaveprocedure$ مسئله ی ما مسئله ای $convex$ است.

ب) همانطور که در صورت سوال گفته شد x^{k+1} جواب بهینه در مجموعه ی نقاط $feasible$ مسئله ی زیر است :

$$\begin{aligned} \min_x \quad & f_0(x) + \hat{g}(x; x^k) \\ \text{s.t.} \quad & f_i(x) \leq 0, \quad i = 1, ..., m \\ & Ax = b \end{aligned}$$

چون جواب بهینه در مجموعه نقاط $feasible$ به ازای هر x $feasible$ دیگری مقدار تابع $objective$ در آن x بزرگتر مساوی مقدار آن در x^{k+1} است. از طرفی می دانیم که برای بسط تیلور مرتبه ی اول دایم که :

$$g(z) \leq \hat{g}(x;z) = g(z) + \nabla g(z)^T(x-z)$$

پس نتیجه می شود که :

$$f_0(x^{k+1}) + g(x^{k+1}) \leq f_0(x^{k+1}) + \hat{g}(x^{k+1};x^k) \leq f_0(x^k) + \hat{g}(x^k;x^k)$$

از طرفی داریم که

$$\hat{g}(x^k;x^k) = g(x^k) + \nabla g(x^k)^T(x^k - x^k) = g(x^k)$$

پس داریم که :

$$f_0(x^{k+1}) + g(x^{k+1}) \leq f_0(x^{k+1}) + \hat{g}(x^{k+1};x^k) \leq f_0(x^k) + g(x^k)$$

در نتیجه حکممان اثبات می شود.

سوال هشتم

الف) دایم که :

$$\begin{aligned} \min_x \quad & x^TAx - 2b^Tx \\ \text{s.t.} \quad & x^TBx = 0 \end{aligned}$$

حالا بجای A و B ماتریس های قطری شده ی آنها را جایگزین می کنیم :

$$\begin{aligned} \min_x \quad & x^TP^Tdiag(\alpha)Px - 2b^Tx \\ \text{s.t.} \quad & x^TP^Tdiag(\beta)Px = 0 \end{aligned}$$

داریم که :

$$x^TP^Tdiag(\alpha)Px = \sum_{i=1}^n (Px)_i\alpha_i(Px)_i = \sum_{i=1}^n (Px)_i^2\alpha_i = (Px \odot Px)^T\alpha = \alpha^Ty$$

حالا داریم که : $Px = sign(Px) \odot \sqrt{y}$ پس خواهیم داشت که :

$$Px = sign(Px) \odot \sqrt{y} \rightarrow x = P^{-1}(sign(Px) \odot \sqrt{y}) \rightarrow -2b^T x =$$

$$-2((sign(Px) \odot \sqrt{y}))^T P^{-T} b = -2\sqrt{y}^T (sign(Px) \odot \gamma)$$

حالا با فرض اینکه علامت حاصل از ضرب $sign(Px)_i$ در γ_i مثبت باشد (اگر علامت x را قرینه کنیم x^TAx، x^TBx تغییری نمی کنند. پس x به این صورت است که علامت Px برابر با علامت $P^{-1}b$ باشد.)خواهیم داشت که عبارت بالا برابر است با : $-2|\gamma|^T\sqrt{y}$ به طور مشابه کاری که برای A کردیم را برای B نیز می توانیم انجام دهیم و به عبارت زیر می رسیم :

$$\begin{aligned} \min_x \quad & \alpha^T y - 2|\gamma|^T \sqrt{y} \\ \text{s.t.} \quad & \beta^T y = 0 \\ & y \succeq 0 \end{aligned}$$

ب)

حل بخش اول سوال ۴.۳۳:

فرض کنیم چنین V وجود دارد آنگاه خواهبم داشت که $\hat{X} + V$ نیمه مثبت معین است و طبق سوال ۹ تمرین اول خواهیم داشت که X و V در یک پایه ی مشترک قطری شدنی اند همچنین با نوشتن رابطه ی X داریم که : $X = Q_1 \Lambda Q_1^T$ پس ماتریس Y ای داریم که : $V = Q_1 Y Q_1^T$ حالا داریم که :

$$tr(A_i V) = tr(A_i Q_1 Y Q_1^T) = tr(Q_1^T A_i Q_1 Y) \rightarrow tr(A_i V) = 0 \leftrightarrow tr(Q_1^T A_i Q_1 Y) = 0$$

از طرفی داریم که :

$$\hat{X} + V \succeq 0 \leftrightarrow Q_1(\Lambda_1 + Y)Q_1^T \succeq 0 \leftrightarrow e^T Q_1 (\Lambda_1 + Y) Q_1^T e \geq 0 \quad for\, all\, e$$

چون ماتریس Q_1 فول رنک است پس عبارت بالا معادل است با :

$$u^T(\Lambda_1 + Y)u \leq 0 \quad for\ all\ u \leftrightarrow \Lambda_1 + Y \succeq 0$$

به طور مشابه برای $\Lambda_1 - Y$ نیز همین رابطه را داریم. پس حکم مسئله اثبات می شود.

حل بخش دوم سوال ۴.۳۳

فرض کنیم که داشته باشیم : $\frac{r(r+1)}{2} > m$ از طرفی تعداد متغییر های آزاد ماتریس

Y برابر است با $\frac{r(r+1)}{2}$ زیرا یک ماتریس متقارن است. با رنک r حالا اگر ماتریس $Q_1 A_i Q_1^T$ را ماتریس B_i در نظر بگیریم داریم که : $tr(B_iV) = 0 = \sum_{j,k}(B_i)_{j,k}Y_{j,k} = 0$ در نتیجه یکی از عناصر V بر حسب سایرین بدست می آید و در نتیجه با m قید m تا از متغییر های آزاد کم می شوند. حالا اگر $m > \frac{r(r+1)}{2}$ برقرار باشد از آنجایی که تعداد عناصر آزاد Y بدون قید $\frac{r(r+1)}{2}$ تا است، در این حالت تعدادی عنصر آزاد باقی می مانند که یعنی رنک ماتریس V بزرگتر از صفر است که در نتیجه جوابی غیر از ماتریس صفر برای V داریم که در شرایط صدق کند پس ماتریس $\hat{X}$ یک $extremepoint$ نیست.

حل بخش سوم سوال ۴.۳۳ در این بخش $relaxation$ را اینگونه در نظر می گیریم که ماتریس X را رنک یک و برابر با xx^T در نظر می گیریم که x یک بردار است. حالا داریم که :

$$tr(A_iX) = tr(A_ixx^T) = x^TA_ix \quad tr(CX) = tr(Cxx^T) = x^TCx$$

پس مسئله ی ما به این صورت بدست می آید که :

$$\begin{aligned} \min_{x} \quad & x^TCx \\ \text{s.t.} \quad & x^TA_ix = b_i, \quad i = 1, ..., m \end{aligned}$$

تحلیل

بخش ب یک تحلیل مهم در باره ی ساختار راه حل ها در $relaxation$ ارائه می دهد که نشان می دهد که اگر تعداد عناصر مستقل در ماتریس V بیشتر از تعداد قیود باشد آنگاه سیستم می تواند چندین راه حل داشته باشد این شرایط نشان می دهد که $\hat{X}$ ممکن است در یک نقطه قابل قبول بزرگتری قرار گیرد که شامل نقاط دیگری نیز می شود که همه ی محدودیت ها را برآورده می کنند و یعنی منحصر به فرد نیست ($exteremepoint$ نیست.) برای اینکه $relaxation$ مربوط به SDP به درستی جواب یا تقریب راه حل $QCQP$ را بگیرد ایده آل است که راه حل بهینه X از SDP داراری رتبه پایین باشد و به طور خاص رنک یک باشد که مستقیما با استفاده از بردار x بدست آمده باشد $X = xx^T$ شرایط بخش ب نشان می دهد که اگر مجموعه ی قابل قبول SDP شامل چند راه حل باشد آنگاه $relaxation$ ممکن است دقیق نباشد یعنی SDP ممکن است راه حل هایی ارائه دهد که مستقیما به راه حل های قابل قبول اصلی $QCQP$ مپ نشوند.
تحت این فرض که مجموعه ی قابل قبول SDP خالی نباشد و محدود باشد و راه حل های بهینه SDP، $extremepoint$ هستند شرط $\frac{r(r+1)}{2} \leq m$ لازم است تا این شرط که $\hat{X}$ یک $extremepoint$ است را برآورده کند این شرط احتمال اینکه جواب های SDP از رنک مینیمال باشند را برآورده می کند. این $relaxation$ را دقیق یا

غیر دقیق می کند و باعث ایجاد یک تناظر بین راه حل های SDP و $QCQP$ می شود.
اگر جواب های SDP، $extremepoint$ نباشند، آنگاه خواهیم داشت که $relaxation$ ما بسیار $loose$ است، به عبارتی ماتریس هایی با رنک بالاتر را تخمین می زند زیرا که از جواب دقیق دور است .
در نهایت برای دقت $relaxation$ مهم است که شرط رنک که در بخش ب گفته شد برقرار باشد تا بعد فضای تعیین شده توسط ماتریس V صفر باشد و $\hat{X}$ یک $extremepoint$ باشد.این باعث ایجاد راه حل هایی می شود که قابل تفسیر اند یا به صورت $QCQP$ باز گردانده می شوند.

مقایسه با مسئله ی خودمان

حالا این مسئله را با سوال بخش الف مقایسه می کنیم داریم که :
در سوال خودمان داشتیم در واقع نسخه ی $relax$ شده ی مسئله ی اصلی را حل می کردیم زیرا که آنجا داریم : $x^TAx, \quad x^TBx$ همچنین یک نرم $-۲b^Tx$ هم به آن اضافه شده است.
پس مسئله ی خودمان یک فرم $relax$ شده از حالت اصلی مسئله ی داده شده در بخش a سوال ۴.۳۳ کتاب اضافه است .همچنین مسئله ی بخش اول این سوال را هم می توانیم به صورت گفته شده ساده کنیم و مسئله ی معادلش که ساده تر است را بنویسیم و سپس آنرا حل کنیم.
