به نام خدا

امتحان پایانترم بهینه‌سازی محدب

کیارش بنی‌هاشم ۹۶۱۰۹۹۶۳

بهار ۱۳۹۹

---

۱. دقت کنید که یک مجموعه نامتناهی است اگر و تنها اگر ماکسیمم قدر مطلق درایه‌هایش نامتناهی باشد. این عملا بدیهی است زیرا اگر تعریف نامتناهی را نامحدود بودن نرم ۲ l۲ بگیریم، چون در فضای nبعدی، نرم ۲ l۲ بزرگتر مساوی ماکسیمم درایه است (طبق فیثاغورث یا حتی نامساوی مثلث) و همچنین از $\sqrt{n}$ برابر نرم l_∞ کمتر است (طبق فیثاغورث)، می‌توان این فرض را کرد.

در نتیجه نامتناهی بودن، معادل با این است که حداقل در یکی از ۲n جهت مختصات، برنامه نامتناهی باشد. معادلا می‌توان برای هر ۲n حالت از بردار $v=(-1)^k e_i$، سعی کنیم که مقدار را بیشینه کنیم. یعنی می‌توان این ۲n برنامه را حل کرد و دید که آیا حداقل یکیشان نامتناهی است یا نه

$$maximize v^T x$$

$$s.t. \quad Ax \leq b \wedge Cx = d$$

اما چون گفته شده ۱ برنامه، همه‌ی این‌ها را با هم قاطی می‌کنیم. یعنی به شکل زیر (در زیر، V همان مجموعه‌ی ۲n برداری است که از در نظر گرفتن خود بردارهای مختصات و منفی‌شان به دست می‌آید.)

$$maximize \sum_{v \in V} v^T x_v$$

$$Ax_v \leq b \wedge Cx_v = d$$

دقت کنید که برای هر کدام از ۲n حالت مختلف، یک متغیر nتایی جدید گرفته‌ایم. ادعا می‌کنیم مثبت بی‌نهایت بودن جواب این برنامه معادل با نامتناهی بودن polyhedron است. اگر polyhedron نامتناهی باشد، یکی از مختصات‌ها به مثبت بی‌نهایت میل می‌کند. در نتیجه x_v متناظر با همه‌ی دیگر را یک نقطه‌ی ثابت می‌گیریم و در نتیجه در تابع هدف، این جملات ثابت می‌مانند. x_v مربوط به این مختصات اما چون می‌تواند به بی‌نهایت میل کند، lp هم می‌تواند به مثبت بی‌نهایت میل کند.

با توجه به این که این راه حل رو میشه تو cvx زد طوری که کار کنه، سراغ این که دوگانش رو دربیاریم و یه جوری ربط بدیم به این که این نامتناهی اگر و تنها اگر اون feasible باشه و اینا نمی‌رم. همچنین چون با همین حالت میشه تو cvx زد، سراغ این که فرم بلوکی ماتریس A' جدید که گنده‌تر شده رو بنویسم هم نمی‌رم.

ابعاد lp اما خب به وضوح خیلی خوب نیست و اگر n تا lp جدا می‌زدیم فکر کنم بهتر میشد چون عملا من هم همین کار رو کردم و در این حالت صرفا امید به cvx و ایناست که خودشون ببینن که A جدید اسپارسه

۲. با استفاده از محاسبه‌ی هسین مساله را حل می‌کنیم. همچنین برای سادگی به‌جای نمادگزاری مساله، از $x^\alpha y^\beta$ استفاده می‌کنیم.

$$\nabla f = \begin{bmatrix} \alpha x^{\alpha-1} y^\beta \\ \beta x^\alpha y^{\beta-1} \end{bmatrix} \implies H = x^{\alpha-2} y^{\beta-2} \begin{bmatrix} \alpha(\alpha-1)y^2 & \alpha\beta xy \\ \alpha\beta xy & \beta(\beta-1)x^2 \end{bmatrix}$$

حال دقت کنید که ضریب پشت ماتریس همواره مثبت است. پس برای بررسی H باید دترمینان آن و نیز ضرایب روی قطر را بررسی کنیم.

(آ) H مثبت‌نیمه‌معین است اگر و تنها اگر

$$\alpha(\alpha-1) \geq 0 \wedge \beta(\beta-1) \geq 0 \wedge x^2y^2\alpha\beta(1-\alpha-\beta) \geq 0$$

در نتیجه اگر $\alpha, \beta \geq 0$ باشند که ممکن نیست چون که باید هر دو حداقل ۱ باشند و در نتیجه نامساوی سوم درست نخواهد بود.

اگر هر دو منفی باشند (منظور نامثبت)، مشکلی نداریم چون که نامساوی‌های برقرارند. پس اگر هر دو منفی باشند تابع محدب است. اگر یکی مثبت و دیگری منفی باشد، مثلا α مثبت باشد، در این صورت

$$\alpha(\alpha-1) \geq 0 \implies \alpha \geq 1, \quad \beta(1-\alpha-\beta) \geq 0 \implies \beta \geq 1-\alpha$$

از طرفی دقت کنید که اگر $\alpha \geq 1, \beta \geq 1-\alpha$، هر ۳ نامساوی برقرارند. پس یک حالت هم این است که $\alpha\beta \leq 0, \alpha+\beta \geq 1$. شکل مورد نظر در ؟؟ آمده است.

(ب) برای دو شرط اول باید داشته باشیم

$$\alpha(\alpha-1) \leq 0 \wedge \beta(\beta-1) \leq 0 \implies 1 \geq \alpha, \beta \geq 0$$

همچنین برای شرط سوم باید داشته باشیم

$$1-\alpha-\beta \geq 0 \implies \alpha+\beta \leq 1$$

پس مقعر بودن معادل است با

$$\alpha, \beta \geq 0 \wedge \alpha + \beta \leq 1$$

شکل مورد نظر در ?? آمده است.

حال به حل بخش امتیازی می‌پردازیم. به این منظور، دقت کنید که هسین روی قطر برابر است با

$$(\prod x_i^{\alpha_i}) \frac{\alpha_i(\alpha_i - 1)}{x_i^2}$$

و خارج از قطر برابر است با

$$(\prod x_i^{\alpha_i}) \frac{\alpha_i \alpha_j}{x_i x_j}$$

که چون $(\prod x_i^{\alpha_i})$ مثبت است، می‌توان آن را در نظر نگرفت. با این توصیفات، به حل مساله می‌پردازیم. **ابتدا مقعر بودن و سپس محدب بودن را بررسی می‌کنیم**

(آ) اولا دقت کنید که اگر $\alpha_i \geq 0$، درایه‌ی iام روی قطر مثبت خواهد بود. در نتیجه قطعا مقعر نیست. پس $\alpha \geq 0$. حال برای هر v دلخواه باید داشته باشیم.

$$\forall v \in R^n : \sum v_i^2 \frac{\alpha_i^2 - \alpha_i}{x_i^2} + 2 \sum_{i \neq j} v_i v_j \frac{\alpha_i \alpha_j}{x_i x_j} \leq 0 \iff (\sum v_i \frac{\alpha_i}{x_i})^2 \leq \sum \alpha_i \frac{v_i^2}{x_i^2} \tag{1}$$

اگر که داشته باشیم $\sum \alpha_i \leq 1$،

$$RHS \geq RHS \times \sum \alpha_i = (\sum \alpha_i)(\sum \alpha_i \frac{v_i^2}{x_i^2})$$

که طبق کوشی شوارتز از LHS بیشتر است و در نتیجه تابع مقعر است. از طرفی اگر 1 برای v دلخواه برقرار باشد، داریم

$$v_i = x_i \implies (\sum \alpha_i)^2 \leq (\sum \alpha_i) \implies \sum \alpha_i \leq 1$$

پس تابع مقعر است اگر و تنها اگر

$$\alpha \geq 0 \wedge \sum \alpha_i \leq 1$$

(ب) فرض کنید که دو تا از α_iها مثبت باشند. در این صورت با در نظر گرفتن تصویر روی همان دو متغیر، باید هنوز تابع محدب بماند. اما با توجه به این که در بخش غیرامتیازی ثابت کردیم که اگر محدب باشد همچین چیزی ممکن نیست، این حالت منتفی است. پس یا همه منفی‌اند و یا دقیقا یکی مثبت است. اگر همه منفی باشند، باید برای هر v دلخواه داشته باشیم (بسط دادن هسین مشابه قبل است)

$$\forall v \in R^n : \sum v_i^2 \frac{\alpha_i^2 - \alpha_i}{x_i^2} + 2 \sum_{i \neq j} v_i v_j \frac{\alpha_i \alpha_j}{x_i x_j} \geq 0 \iff (\sum v_i \frac{\alpha_i}{x_i})^2 \geq \sum \alpha_i \frac{v_i^2}{x_i^2} \tag{2}$$

که همواره برقرار است زیرا سمت راست نامثبت است و سمت چپ نامنفی. اگر هم دقیقا یکی از α_iها مثبت باشد و بقیه نامثبت، با قرار دادن $v_i = x_i$ نتیجه می‌گیریم که

$$(\sum \alpha_i)^2 \geq \sum \alpha_i$$

و در نتیجه یا $\sum \alpha_i \leq 0$، یا $\sum \alpha_i \geq 1$. در حالت اول، اگر همه‌ی x_i هایی که ضریبشان منفی است را یکی کنیم (یعنی یک متغیر کنیم. این کار ممکن است چون precomposition با آفین)، تابع باید محدب بماند. اما خب محدب نمی‌ماند زیرا الان یک α_i مثبت داریم، یک α_i منفی و جمعشان هم منفی است که در شرایط بخش غیرامتیازی صدق نمی‌کند چون در حالت یکی مثبت، یکی منفی، باید جمع حداقل 1 می‌شد.
پس در حالت دوم هستیم و در نتیجه

$$\sum \alpha_i \geq 1$$

ثابت می‌کنیم که همین شرایط کافیست. به این منظور ابتدا حالتی که $\sum \alpha_i = 1$ است را بررسی می‌کنیم. در این حالت، اگر فرض کنیم که $\alpha_n > 0$ و بقیه نامثبت‌اند، می‌دانیم $x \in R^{n-1} \to \prod x_i^{\alpha_i}$ محدب است. در نتیجه طبق قانون پرسپکتیو، تابع زیر محدب است

$$x \in R^{n-1}, g(t, x) = t \prod ((\frac{x_i}{t})^{\alpha_i}) = t^{1-\sum \alpha_i} \prod x_i^{\alpha_i}$$

که همان تابع ماست با این تفاوت که به‌جای x_n از t استفاده کرده‌ایم. پس این حالت حل شد. در حالتی که $\sum \alpha_i > 1$ هم می‌توانیم تعریف کنیم

$g(x) = \prod x_i^{\frac{\alpha_i}{\sum \alpha_i}}$. طبق چیزی که الان ثابت کردیم، g محدب است. پس چون تابع زیر محدب است (چون ۱ $\geq \sum \alpha_i$) و همچنین صعودی است

$$x \in R^+, h(x) = x^{\sum \alpha_i}$$

نتیجه می‌گیریم که تابع اصلی هم محدب است. پس برای این بخش هم ۲ حالت وجود دارد. حالت اول این که

$$\alpha \leq ۰$$

و حالت دوم هم این که دقیقا یکی از α_iها مثبت است و همچنین

$$\sum \alpha_i \geq ۱$$

۴. (آ) صورت حکم چیزی به فرم فلان اگر و فقط اگر بهمان است. من ثابت می‌کنم که بهمان اگر و تنها اگر فلان. دقت کنید که همان مطلب ثابت شده است و صرفا چون تشکیل دوگان از این سمت برایم راحت‌تر بود این کار را می‌کنم.
مساله‌ی زیر را در نظر بگیرید

$$minmize_\mu \quad -c^T\mu$$

$$s.t. \quad S^T\mu \leq ۰$$

این مساله به وضوح محدب است (اصلا آفین است). همچنین شرط slater برقرار است زیرا اگر قرار دهیم $\mu = ۰$، در نامساوی‌ها (که همه خطی‌اند) صدق می‌کند. در نتیجه strong dualtiy برایش برقرار است. از طرفی دقت کنید که جواب این مساله بیشتر مساوی ۰ است، اگر و تنها اگر بخش بهمان سوال برقرار باشد یعنی برای هر $S^T\mu \leq ۰$ بتوان گفت که $c^T\mu \leq ۰$. از طرفی با تشکیل دوگان داریم

$$\mathcal{L} = -c^T\mu + v^TS^T\mu = (-c + Sv)^T\mu$$

که با اینفیمم‌گیری روی μ، منفی بی‌نهایت است اگر ضریب ناصفر باشد و در غیر این صورت ۰ است. پس مساله‌ی دوگان به فرم زیر است

$$maximize \quad ۰$$

$$s.t. \quad Sv - c = ۰ \wedge v \geq ۰$$

از طرفی جواب این مساله برابر با ۰ است، اگر و تنها اگر feasible باشد که معادل است با این که بخش فلان مساله درست باشد یعنی c یک بردار thermodinamically feasible باشد. پس حکم ثابت شد چون ثابت کردیم که فلان معادل است با ۰ بودن جواب مساله‌ی دوگان که معادل است با ۰ بودن مساله‌ی اصلی که معادل است با بهمان.

(ب) فرض کنید که حکم غلط باشد. مساله‌ی زیر را در نظر بگیرید

$$minimize \quad -t$$

$$s.t. \quad S^Tm = ۰ \wedge m \geq t$$

در این صورت جواب این مساله، قطعا مقداری منفی‌ای نخواهد بود زیرا اگر منفی باشد، t مثبت است و در نتیجه $m > ۰$ و همچنین $m^TS = ۰$ که حکم است و در نتیجه با فرض خلف در تناقض است. پس جواب مساله بزرگتر مساوی ۰ است. از طرفی اگر قرار دهیم $m = ۰ \wedge t = ۰$، به وضوح به ۰ می‌رسیم. پس جواب این مساله ۰ است.
مساله به وضوح محدب است (مجددا همه‌چی آفین است) و شرط slater برقرار است چون می‌توانیم قرار دهیم $t = -۱, m = ۰$ (البته اگر $t = ۰$ قرار می‌دادیم هم اوکی بود). در نتیجه strong dualtiy داریم. مساله‌ی دوگان را تشکیل می دهیم.

$$\mathcal{L} = -t + u^T(S^Tm) + \lambda^T(t۱ - m) = t(\lambda^T۱ - ۱) + (Su - \lambda)^Tm$$

که با اینفیمم‌گیری روی t, m، برابر با منفی بی‌نهایت است اگر یکی از ضرایب ناصفر باشد و در غیر این صورت ۰ است. پس مساله‌ی دوگان به فرم زیر است.

$$maximize \quad ۰$$

$$s.t. \quad \lambda \geq ۰ \wedge Su = \lambda \wedge \sum ۱^T\lambda = ۱$$

چون جواب مساله‌ی اصلی ۰ بود، جواب این مساله هم ۰ است و در نتیجه feasible است. پس چون داریم

$$Su = \lambda \geq 0$$

طبق فرض

$$Su = 0 \implies \lambda = 0 \implies 1 = 1^T\lambda = 0$$

که به وضوح تناقض است.

۵. (آ) بیانی که از مساله در حال حاضر داریم این است که قیدها به شرح زیراند

$$1 \leq \frac{l}{w} \leq 2, \quad 10 \leq w \leq 20, \quad 20 \leq l \leq 30, \quad lw \geq 300$$

و با توجه به این که مساحت فیلتر برابر است با lw و محیط برابر است با $2l + 2\pi\frac{w}{2}$، و ارزش مساحت دوبرابر محیط است، تابع زیرا را می‌خواهیم که کمینه کنیم.

$$2lw + 2l + 2\pi\frac{w}{2}$$

اولا دقت کنید که قید $w \geq 10$ اضافی است زیرا از $\frac{l}{w} \leq 2$ و $l \geq 20$ نتیجه می‌شود. تعریف کنید.

$$s = \sqrt{lw}$$

دقت کنید که طبق قیدهای دیگر می‌توان فرض کرد کلا l, w مثبتند. در نتیجه

$$s^2 = lw \implies w = \frac{s^2}{l}, \quad \frac{l}{w} = \frac{l^2}{lw} = \frac{l^2}{s^2}$$

در نتیجه قید روی $\frac{l}{w}$ را می‌توان به فرم زیر نوشت

$$1 \leq \frac{l^2}{s^2} \leq 2 \iff 1 \leq \frac{l}{s} \leq \sqrt{2} \iff s \leq l \leq s\sqrt{2}$$

قید روی w هم می‌توان به فرم زیر نوشت

$$w \leq 20 \iff \frac{s^2}{l} \leq 20$$

که محدب است.

قید روی l هم که پابرجاست

$$20 \leq l \leq 30$$

قید روی مساحت هم به فرم زیر قابل بیان است.

$$s^2 \geq 300 \iff s \geq \sqrt{300}$$

با این تغییر متغیرها، تابعی که باید کمینه کنیم هم به فرم زیر است.

$$2s^2 + 2l + \pi\frac{s^2}{w}$$

(ب) جواب ۳۰۰ است. این کار را می‌توان با $cvxpy$ دید که کدش آمده است. همچنین می‌توان بدون آن دید (صرفا به امید نمره‌ی امتیازی و اینا نوشتم این توضیح رو). دقت کنید که اگر قید مساحت را در نظر نگیریم، طبیعتا هر چه l, w کمتر شوند بهتر است. پس $l = 20, w = 10$ جواب بهینه را می‌دهد. اما مشخص است که در این حالت، مساحت کافی نخواهد بود. حال که قید مساحت را اضافه می‌کنیم، اگر tight نباشد، کاملا احمقانه است زیرا معنی‌اش این است که ضریب لاگرانژ مرتبط با این قید ۰ است و در نتیجه لاگرانژین مساله، لاگرانژین مساله‌ی بدون قید مساحت است. در نتیجه چون این‌جا strong duality داریم (کلا مساله‌ها هم محدب‌اند و هم شرطها رو میشه اکید برقرار کرد. این رو دیگه چک نکردم ولی خب قطعا میشه)، الان جواب مساله‌ی جدید همه‌ی شرطهای KKT رو برای مساله‌ی قبلی هم خواهد داشت و در نتیجه برای اون هم بهینه است که خب گفتیم نمیشه چون تو مساله‌ی قبلی حتما مساحت میشد ۲۰۰.

۶. (آ) با توجه به قیدها، مشخص است که $l, w > 0$. با این حساب، تعریف کنید

$$l' = ln(l), w' = ln(w)$$

قیدهای روی $l, w, \frac{l}{w}$ به فرم زیر درمی‌آیند.

$$\circ \leq l' - w' \leq ln(۲), \quad ln(۱\circ) \leq w' \leq ln(۲\circ), \quad ln(۲\circ) \leq l' \leq ln(۳\circ)$$

همچنین قید روی مساحت به شکل زیر است

$$l' + w' \geq ln(۳\circ\circ)$$

برای محاسبه‌ی هزینه هم داریم

$$cost = ۲lw + ۲l + ۲\pi\frac{w}{۲} = ۲e^{l'+w'} + ۲e^{l'} + \pi e^{w'}$$

۷. (آ) اگر D و K را با بردار جایگزین کنیم (با توجه به این که قطری هستند)، داریم $A_{i,:} = d_i + k_iG_{i,:}$. همچنین

$$\gamma tr(D^{-۱}) + tr(K^{-۱}) = \sum \frac{\gamma}{d_i} + \sum \frac{۱}{k_i}$$

با این توصیفات مساله به فرم زیر در می‌آید.

$$minimize \quad \lambda_{max}(A)$$
$$s.t. \quad d \geq \circ \wedge k \geq \circ$$
$$A_{i,:} = d_i + k_iG_{i,:}$$
$$\sum \frac{\gamma}{d_i} + \sum \frac{۱}{k_i} \leq u$$

که با توجه به محدب بودن بزرگترین مقدار ویژه و نیز $\frac{۱}{x}$، مساله به فرم مناسب درآمد.

(ب)

۸. فرض کنید که می‌خواهیم بررسی کنیم که آیا n نفر جا می‌شوند یا نه. باید مساله‌ی زیر را بررسی کنیم که feasible است یا نه.

$$\circ \leq x_i \leq l, \quad \circ \leq y_i \leq w$$

$$||x_i - x_j, y_i - y_j||_p \geq ۲ \quad i < j$$

حال دقت کنید که

$$||a,b||_۱ \geq k \iff a + b \geq k \vee a - b \geq k \vee -a + b \geq k \vee -a - b \geq k$$

و به طور مشابه

$$||a,b||_\infty \geq k \iff a \geq k \vee b \geq k \vee -a \geq k \vee -b \geq k$$

نکته‌ی مهم این است که اگر علامت a و b را بدانیم، می‌توانیم قید بالا را به فرم آفین بنویسیم زیرا برای این که بررسی کنیم که آیا حداقل یکی از عبارات بزرگتر یا مساوی k است، کافیست آن عبارتی که در آن علامت‌ها مثبتند را بررسی کنیم چون بیشترین است.
با این توصیفات، برای حل مساله‌ی feasiblity برای n نفر، فرض می‌کنیم که بر اساس x مرتب‌اند به طوری که $x_i \leq x_{i+۱}$. برای بررسی ترتیب yها هم، همه‌ی $n!$ حالت ممکن را در نظر می‌گیریم. یعنی روی همه‌ی این حالات فور می‌زنیم. در هر کدام از این حالات، چون علامت را می‌دانیم، می‌توانیم برنامه‌ی محدب را تشکیل دهیم و بررسی کنیم که آیا n نفر جا می‌شوند یا خیر. حال اگر جواب نهایی T باشد، می‌توانیم با شروع از ۱ و امتحان همه‌ی اعداد تا زمانی که به T برسیم فور بزنیم [۱] پس در کل

$$\sum_{i=۱}^{T} i!$$

تا باید برنامه حل کنیم. عبارت فوق از نظر مرتبه عملا $T!$ است زیرا

$$\sum_{i=۱}^{T} i! = \sum_{i=۱}^{T-۱} i! + T! \leq (T-۱) \times (T-۱)! + T! \leq ۲T!$$

کران فوق برای هر دو بخش الف و ب برقرار است. در صورت نیاز، می‌توان کران‌هایی بر حسب l, w هم داد. به طور دقیق‌تر، در حالتی که $p = ۱$، اگر حول هر نفر و به مرکزیت آن یک دایره‌ی l_∞ به شعاع ۱ بزنیم (در واقع همان مربع خودمان هستند)، این دوایر همدیگر را قطع نمی‌کنند زیرا اگر قطع کنند یعنی دو نقطه نزدیک‌تر از ۲ فاصله دارند. همه‌ی دوایر در مستطیل به طول و عرض $l + ۲, w + ۲$ جا می‌شوند و همچنین مساحت هرکدام حداقل ۴ است.

[۱] چون حل مساله برای T به مراتب سخت‌تر از حل آن برای $T - ۱$ است، باینری سرچ زدن کار اشتباهی است.

در نتیجه

$$p = \infty \implies ۴T \leq (l+۲)(w+۲) \implies T \leq \frac{(l+۲)(w+۲)}{۴}$$

برای نمر l۱ هم به طور مشابه همه‌چیز در همان مستطیل جا می‌شوند اما مربع‌ها مساحتشان نصف شده پس کران بالا به شکل

$$T \leq \frac{(l+۲)(w+۲)}{۲}$$

در می‌آید. کد این را من نزدم چون فکر کنم در حالتی که در هر ثانیه ۱۰ به توان ۹ تا برنامه حل شوند، ۷۷ سال طول می‌کشد که تمام شود!