به نام خدا

بهینه‌سازی محدّب ۱ (۱-۲۵۷۵۶)

تمرین سری اوّل

ترم بهار ۱۴۰۱-۰۲

دانشکده‌ی مهندسی برق

دانشگاه صنعتی شریف

استاد: دکتر محمّدحسین یاسائی میبدی

مهلت تحویل: جمعه ۱۹ اسفند، ساعت ۲۳:۵۹

$(*)$ مسائلی که با ستاره مشخّص شده‌اند امتیازی هستند و حل‌کردن آن‌ها نمره‌ی امتیازی خواهد داشت!

۱

۱. بردار $\mathbf{x} = [x_1, x_2, \ldots, x_n]^\top \in \mathbb{R}^n$ را در نظر بگیرید. ماتریس $\mathbf{P} \in \mathbb{R}^{n\times n}$ را با فرض $\mathbf{P}^\top = \mathbf{P}$ به گونه‌ای بیابید که داشته باشیم:

$$\mathbf{x}^\top \mathbf{P} \mathbf{x} = \sum_{i=1}^{n-1} (x_{i+1} - x_i)^2.$$

۲. به ازای ماتریس $\mathbf{P}$ که در بخش قبل محاسبه کردید، نشان دهید $\mathbf{P} \succeq 0$. آیا می‌توان گفت که $\mathbf{P} \succ 0$ است؟

۲

۱. به ازای هر ماتریس $\mathbf{A} \in \mathbb{R}^{n\times n}$ نشان دهید اگر برداری مانند $\mathbf{x} \in \mathbb{R}^n$ وجود داشته باشد، به گونه‌ای که $\mathbf{A}\mathbf{x} = 0$، آن‌گاه داریم $\det\{\mathbf{A}\} = 0$.

۲. به ازای هر ماتریس $\mathbf{A} \in \mathbb{R}^{n\times n}$ نشان دهید شرط کافی برای آن‌که $\mathbf{I} + \mathbf{A}$ وارون‌پذیر باشد، آنست که $\|\mathbf{A}\|_2 < 1$ باشد.

۳

درستی یا نادرستی موارد زیر را با ذکر دلیل مشخص کنید.

۱. اگر $\mathbf{A} \in \mathbb{R}^{m\times n}$ و $\mathbf{B} \in \mathbb{R}^{n\times p}$ ماتریس هایی با رتبه‌ی کامل باشند و $\mathbf{AB} = 0$ باشد، آن‌گاه داریم: $n \geq m + p$.

۲. اگر به ازای برخی مقادیر صحیح $k \geq 1$ داشته باشیم $\mathbf{A}^k = 0$، آنگاه $\mathbf{A} - \mathbf{I}$ ماتریسی با رنک کامل است.

۳. به ازای هر $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n}$ بردارهای ویژه‌ی $\mathbf{AB}$ برابر با بردارهای ویژه‌ی $\mathbf{BA}$ هستند.

۴

نرم فروبینیوس[1] برای ماتریس $\mathbf{A} \in \mathbb{R}^{m\times n}$ به صورت زیر تعریف می‌شود:

$$\|\mathbf{A}\|_F = \sqrt{\mathrm{Tr}\left\{\mathbf{A}^\top \mathbf{A}\right\}}$$

۱. ثابت کنید:

$$\|\mathbf{A}\|_F = \sqrt{\sum_{i=۱}^{m}\sum_{j=۱}^{n} |A_{i,j}|^۲}.$$

۲. اگر مقادیر تکین ماتریس $\mathbf{A}$ را با $\sigma_۱, \sigma_۲, \ldots, \sigma_r$ نشان دهیم، ثابت کنید:

$$\|\mathbf{A}\|_F = \sqrt{\sum_{i=۱}^{r} \sigma_i^۲}.$$

۳. ثابت کنید:

$$\sigma_{\max}(\mathbf{A}) \leq \|\mathbf{A}\|_F \leq \sqrt{r}\sigma_{\max}(A)$$

۵

۱. ماتریس‌های $\mathbf{A} \in \mathbb{R}^{n\times n}$، $\mathbf{B} \in \mathbb{R}^{n\times k}$، $\mathbf{C} \in \mathbb{R}^{k\times n}$ و $\mathbf{D} \in \mathbb{R}^{k\times k}$ را در نظر بگیرید. فرض کنید ماتریس‌های $\mathbf{A}, \mathbf{D}$ معکوس‌پذیر باشند. بدون استفاده از تعریف ماتریس وارون برای ماتریس $\mathbf{A} - \mathbf{B}\mathbf{D}^{-۱}\mathbf{C}$ نشان دهید:

$$(\mathbf{A} - \mathbf{B}\mathbf{D}^{-۱}\mathbf{C})^{-۱} = \mathbf{A}^{-۱} + \mathbf{A}^{-۱}\mathbf{B}(\mathbf{D} - \mathbf{C}\mathbf{A}^{-۱}\mathbf{B})^{-۱}\mathbf{C}\mathbf{A}^{-۱}.$$

راهنمایی: سعی کنید معکوس‌پذیری ماتریس فوق را با تعریف یک متغیّر کمکی به یک دستگاه معادلات با ماتریس ضرایب $\begin{bmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{C} & \mathbf{D}\end{bmatrix}$ تبدیل کنید.

۲. شرایط معکوس‌پذیری و معکوس ماتریس $\mathbf{I} + \mathbf{u}\mathbf{v}^\top$ را به دست آورید که در آن $\mathbf{I}$ ماتریس همانی است و داریم $\mathbf{I} \in \mathbb{R}^{n\times n}$ و $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$.

۳. به ازای هر ماتریس معکوس‌پذیر $\mathbf{A} \in \mathbb{R}^{n\times n}$ و هر دو بردار $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ نشان دهید:

$$(\mathbf{A} + \mathbf{u}\mathbf{v}^\top)^{-۱} = \mathbf{A}^{-۱} - \frac{\mathbf{A}^{-۱}\mathbf{u}\mathbf{v}^\top\mathbf{A}^{-۱}}{۱ + \mathbf{v}^\top\mathbf{A}^{-۱}\mathbf{u}}.$$

۴. یکی از مطالب پرکاربردی که در طول درس ظاهر می‌شود، بررسی مثبت معین یا مثبت نیمه معین بودن یک ماتریس متقارن است. به علاوه بسیاری از اوقات ما به کمک روش‌های ریاضی می‌توانیم قیود غیرخطّی در مسائل بهینه‌سازی را به قیود خطّی تبدیل کنیم تا مسئله به فرم یک مسئله‌ی بهینه‌سازی محدب در بیاید. یکی از روش‌هایی که در این موضوع می‌تواند کمک کننده باشد استفاده از مفهومی به نام Schur Complement در ماتریس‌های متقارن است.

ماتریس متقارن دلخواه $\mathbf{M} \in \mathbb{R}^{n\times n}$ را به صورت زیر در نظر بگیرید:

$$\mathbf{M} = \begin{bmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{B}^\top & \mathbf{C}\end{bmatrix},$$

که در آن $\mathbf{A} \in \mathbb{R}^{r\times r}$،

$\mathbf{B} \in \mathbb{R}^{r\times(n-r)}$ و $\mathbf{C} \in \mathbb{R}^{(n-r)\times(n-r)}$. همچنین فرض کنید $\mathbf{C}$ معکوس‌پذیر باشد. ماتریس $\mathbf{A} - \mathbf{B}^\top\mathbf{C}^{-۱}\mathbf{B}$ را ماتریس Schur Complement $\mathbf{C}$ در $\mathbf{M}$ می‌نامیم. نشان دهید:

(آ) $\mathbf{M} \succ ۰$ اگر و تنها اگر $\mathbf{C} \succ ۰$ و $\mathbf{A} - \mathbf{B}\mathbf{C}^{-۱}\mathbf{B}^\top \succeq ۰$.

(ب) اگر $\mathbf{C} \succ ۰$، آنگاه $\mathbf{M} \succeq ۰$ اگر و تنها اگر $\mathbf{A} - \mathbf{B}\mathbf{C}^{-۱}\mathbf{B}^\top \succeq ۰$.

[1] Frobenious Norm

۶

در این سوال قصد داریم نامساوی ون-نیومن را اثبات نماییم. این نامساوی می‌گوید فرض کنید $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times n}$ با مقادیر تکین $\sigma_1(\mathbf{A}) \geq \ldots \geq \sigma_r(\mathbf{A})$ و $\sigma_1(\mathbf{B}) \geq \ldots \geq \sigma_r(\mathbf{B})$ باشند که $r = \min\{m, n\}$ ، آن‌گاه:

$$|\mathrm{Tr}\left\{\mathbf{A}^\top \mathbf{B}\right\}| \leq \sum_{i=1}^{r} \sigma_i(\mathbf{A})\sigma_i(\mathbf{B}).$$

۱. در ابتدا نشان دهید که اگر $\mathbf{U}$ و $\mathbf{V}$ دو ماتریس مثبت نیمه معین هرمیتی باشند و λ_1 متناظر با بزرگ‌ترین مقدار ویژه‌ی یک ماتریس باشد، نشان دهید که:

$$\lambda_1(\mathbf{UV}) \leq \lambda_1(\mathbf{U})\lambda_1(\mathbf{V}).$$

۲. حال به کمک بخش قبل نشان دهید که:

$$\prod_{i=1}^{k} \lambda_i(\mathbf{UV}) \leq \prod_{i=1}^{k} \lambda_i(\mathbf{U}) \prod_{i=1}^{k} \lambda_i(\mathbf{V}).$$

که در آن $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_k$ مقادیر ویژه‌ی ماتریس‌های متناظر به ترتیب نزولی هستند.

۳. می‌دانیم که مقادیر تکین یک ماتریس نوعی مانند $\mathbf{A}$ در رابطه‌ی زیر صدق می‌کنند.

$$\sigma_i(\mathbf{A}) = (\lambda_i(\mathbf{AA}^\dagger))^{\frac{1}{2}}$$

حال به کمک رابطه‌ی بالا و قضایای ثابت شده در بخش قبل، قضیه ون-نیومن را ثابت نمایید.

۷

در این سؤال می‌خواهیم یکی از کاربردهای مهم تجزیه مقادیر تکین برای ماتریس‌ها را بیان کنیم و آن این است که به تعبیری بهترین تخمین رنک پایین از یک ماتریس را می‌توان از آن‌ها به دست آورد.
فرض کنید $\mathbf{A} \in \mathbb{R}^{m\times n}$ با تجزیه‌ی مقادیر تکین به صورت $\mathbf{A} = \mathbf{U\Sigma V}^\top$ وجود داشته باشد. برای یک عدد صحیح $k \leq \min\{m, n\}$ ماتریس $\mathbf{A}_k \in \mathbb{R}^{m\times n}$ را که قرار است تقریبی از ماتریس $\mathbf{A}$ باشد، به صورت $\mathbf{A}_k = \mathbf{U}_k \mathbf{\Sigma}_k \mathbf{V}_k^\top$ تعریف می‌کنیم که $\mathbf{U}_k$ و $\mathbf{V}_k$ به ترتیب k ستون اول $\mathbf{U}$ و $\mathbf{V}$ و همچنین $\mathbf{\Sigma}_k$ بلوک $k \times k$ سمت بالا و چپ از ماتریس $\mathbf{\Sigma}$ است (که شامل k بزرگ‌ترین مقادیر تکین ماتریس $\mathbf{A}$ می‌شود). به تعبیر دیگر شما می‌توانید $\mathbf{A}_k$ را به صورت $\mathbf{A}_k = \sum_{i=1}^{k} \sigma_i u_i v_i^\top$ نیز بنویسید که در آن $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{\min\{m,n\}}$ مقادیر تکین ماتریس $\mathbf{A}$ هستند.

۱. برای به دست آوردن یک تقریب خوب از ماتریس $\mathbf{A}$ مسئله‌ی بهینه‌سازی زیر را در نظر بگیرید:

$$\min_{\mathbf{X}:\mathrm{rank}(\mathbf{X})\leq k} ||\mathbf{A} - \mathbf{X}||_F.$$

نشان دهید $\mathbf{A}_k$ جواب مسئله‌ی بهینه‌سازی فوق است.

۲. حال به جای نرم فروبینوس در مسئله‌ی بهینه‌سازی قبل، نرم عملگری یک ماتریس را در نظر می‌گیریم. یعنی مسئله‌ی بهینه‌سازی قسمت قبل را به صورت زیر تغییر می‌دهیم:

$$\min_{\mathbf{X}:\mathrm{rank}(\mathbf{X})\leq k} ||\mathbf{A} - \mathbf{X}||_2.$$

نشان دهید هم‌چنان جواب مسئله‌ی بهینه‌سازی فوق $\mathbf{A}_k$ است.

راهنمایی: $k + 1$ ستون اول ماتریس $\mathbf{V}$ را در نظر بگیرید و به کمک ترکیب خطی آن‌ها، یک بردار در فضای پوچ ماتریس $\mathbf{X}$ که متغیر مسئله‌ی بهینه‌سازی فوق است، بسازید.