به نام خدا

بهینه‌سازی محدب ۱ (۲۵۷۵۶)

تمرین شماره ۱

نیم‌سال اول ۱۴۰۴-۱۴۰۳

۱- نابرابری *Von Neumann* و نُرم اتمی

الف) برای ماتریس‌های متقارن $\mathbf{A}, \mathbf{B} \in \mathbb{S}^n$ نابرابری زیر را اثبات کرده و یک شرط کافی برای حالت تساوی آن بیان نمایید.

$$\mathrm{tr}\{\mathbf{AB}\} \leq \sum_{i=1}^{n} \lambda_i(\mathbf{A})\lambda_i(\mathbf{B})$$

ب) برای ماتریس‌های غیر مربعی $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times n}$ نابرابری زیر (که نسخه‌ای تعمیم‌یافته از نابرابری قسمت الف است) را اثبات کرده و یک شرط کافی تساوی برای آن بیان کنید. (مقادیر تکین σ_i را به ترتیب نزولی در نظر بگیرید.)

$$\mathrm{tr}\{\mathbf{A}^T\mathbf{B}\} \leq \sum_{i=1}^{\min\{m,n\}} \sigma_i(\mathbf{A})\sigma_i(\mathbf{B})$$

ج) نُرم ۲ عملگری ماتریس ($\|\mathbf{A}\|_2 = \sigma_1(\mathbf{A})$) را در نظر گرفته و با استفاده از نابرابری قسمت ب، نشان دهید که نُرم دوگان آن برابر نُرم اتمی است؛ یعنی خواهیم داشت:

$$\|\mathbf{B}\|_* = \sup\{\mathrm{tr}\{\mathbf{A}^T\mathbf{B}\} \mid \|\mathbf{A}\|_2 \leq 1\} = \sum_{i=1}^{\min\{m,n\}} \sigma_i(\mathbf{B})$$

د) بررسی کنید که نُرم اتمی چه کاربردی در بهینه‌سازی دارد؟

۲-مساله حداقل نُرم اقلیدسی

دستگاه معادلات $\mathbf{Ax} = \mathbf{b}$ را در نظر بگیرید که در آن $\mathbf{A} \in \mathbb{R}^{m\times n}$ یک ماتریس کشیده افقی ($m < n$) با رتبه کامل سطری است. بردار $\mathbf{x}^* = \mathbf{A}^T(\mathbf{AA}^T)^{-1}\mathbf{b}$ را در نظر بگیرید. مشخصاً این بردار در معادله $\mathbf{Ax} = \mathbf{b}$ صدق می‌کند و لذا یک جواب این دستگاه است. در این تمرین قصد داریم با استفاده از تعامد، نشان دهیم که $\mathbf{x}^*$ در میان جواب‌های دستگاه کمترین نُرم اقلیدسی را دارد.

الف) نشان دهید که برای هر $\mathbf{x}$ دلخواه از مجموعه جواب دستگاه، $(\mathbf{x} - \mathbf{x}^*) \in N(\mathbf{A})$ است.

ب) با استفاده از نتیجه قسمت الف، نشان دهید که برای هر $\mathbf{x}$ دلخواه از مجموعه جواب دستگاه، $(\mathbf{x} - \mathbf{x}^*) \perp \mathbf{x}^*$ است. با توجه به تئوری اساسی جبرخطی، $\mathbf{x}^*$ متعلق به کدام زیرفضای اساسی ماتریس $\mathbf{A}$ است.

ج) با استفاده از نتیجه قسمت ب، نشان دهید که $\|\mathbf{x}\|_2^2 \geq \|\mathbf{x}^*\|_2^2$ است. به عبارتی، $\mathbf{x}^*$ از بین تمامی پاسخ‌های دستگاه کمترین نُرم اقلیدسی را دارد.

د) ماتریس نگاشت به فضای سطری و فضای پوچی را به دست آورید. توجه کنید که این ماتریس‌ها قرار است بردار $\mathbf{x}$ را به زیرفضاهای گفته‌شده نگاشت کنند و سایز آن‌ها $n \times n$ است. رتبه این ماتریس‌ها را نیز بیان کنید.

۳- ماتریس‌های بلوکی

الف) ماتریس بلوکی $\mathbf{M} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}$ را در نظر بگیرید که در آن $\mathbf{A} \in \mathbb{R}^{n\times n}$، $\mathbf{B} \in \mathbb{R}^{n\times m}$، $\mathbf{C} \in \mathbb{R}^{m\times n}$ و $\mathbf{D} \in \mathbb{R}^{m\times m}$ است. با نوشتن رابطه وارون ماتریس بلوکی، اتحاد زیر که به لم معکوس ماتریس موسوم است را اثبات کنید. فرض کنید که معکوس ماتریس‌های موجود در این رابطه وجود داشته باشند.

$$(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1} = \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{B}(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{C}\mathbf{A}^{-1}$$

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

ج) 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:

$$(\mathbf{A} + \mathbf{u}\mathbf{v}^T)^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{u}\mathbf{v}^T\mathbf{A}^{-1}}{1 + \mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}}$$

۴- مکمل شور

یکی از مفاهیم پرکاربرد در درس، بررسی مثبت معین یا نیمه‌معین بودن یک ماتریس متقارن است. ماتریس بلوکی متقارن $\mathbf{M} \in \mathbb{R}^{n\times n}$ را به صورت $\mathbf{M} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^T & \mathbf{C} \end{bmatrix}$ در نظر بگیرید که در آن $\mathbf{A} \in \mathbb{R}^{k\times k}$، $\mathbf{B} \in \mathbb{R}^{k\times(n-k)}$ و $\mathbf{C} \in \mathbb{R}^{(n-k)\times(n-k)}$ است. با فرض معکوس‌پذیری $\mathbf{C}$، ماتریس $\mathbf{S} = \mathbf{C} - \mathbf{B}^T\mathbf{A}^{-1}\mathbf{B}$ را مکمل شور ماتریس $\mathbf{A}$ در $\mathbf{M}$ می‌نامیم. نشان دهید:

الف) $\mathbf{M} \succ 0$ است اگر و تنها اگر $\mathbf{A} \succ 0$ و $\mathbf{S} \succ 0$ باشد.

ب) اگر $\mathbf{A} \succ 0$ باشد، آنگاه $\mathbf{M} \succcurlyeq 0$ است اگر و تنها اگر $\mathbf{S} \succcurlyeq 0$ باشد.

ج) $\mathbf{M} \succcurlyeq 0$ است اگر و تنها اگر $\mathbf{A} \succcurlyeq 0$، $\mathbf{S} \succcurlyeq 0$ و $\mathbf{B}^T(\mathbf{I} - \mathbf{A}\mathbf{A}^{\dagger}) = 0$ باشد.

د) نشان دهید ماتریس بلوکی $\mathbf{M}$ مثبت معین است اگر و فقط اگر $\mathbf{A}$ و $\mathbf{C}$ مثبت معین بوده و تمامی مقادیر تکین ماتریس $\mathbf{A}^{-\frac{1}{2}}\mathbf{B}\mathbf{C}^{-\frac{1}{2}}$ کوچکتر از ۱ باشند.

۵- مقدار ویژه و تکین

الف) نشان دهید در صورتی‌که ماتریس‌های $\mathbf{A} \in \mathbb{M}_m(\mathbb{C})$ و $\mathbf{B} \in \mathbb{M}_n(\mathbb{C})$ مقادیر ویژه مشترک نداشته باشند، معادله $\mathbf{AX} = \mathbf{XB}$ تنها جواب بدیهی $\mathbf{X} = \mathbf{0}$ را دارد.

ب) ماتریس متقارن $\mathbf{A}$ را با مقادیر ویژه $\lambda_1 \geq \cdots \geq \lambda_n$ در نظر بگیرید. نشان دهید:

$$\lambda_k = \max_{\substack{S\subseteq\mathbb{R}^n \\ \dim(S)=k}} \min_{\substack{\mathbf{x}\in S \\ \mathbf{x}\neq\mathbf{0}}} \frac{\mathbf{x}^T\mathbf{A}\mathbf{x}}{\mathbf{x}^T\mathbf{x}} = \min_{\substack{C\subseteq\mathbb{R}^n \\ \dim(C)=n-k+1}} \max_{\substack{\mathbf{x}\in C \\ \mathbf{x}\neq\mathbf{0}}} \frac{\mathbf{x}^T\mathbf{A}\mathbf{x}}{\mathbf{x}^T\mathbf{x}}$$

همچنین با استفاده از نتیجه فوق، برای ماتریس $\mathbf{A} \in \mathbb{M}_{m\times n}(\mathbb{R})$، حاصل عبارت $\min\limits_{\substack{C\subseteq\mathbb{R}^n \\ \dim(C)=n-k+1}} \max\limits_{\substack{\mathbf{x}\in C \\ \mathbf{x}\neq\mathbf{0}}} \frac{\|\mathbf{A}\mathbf{x}\|_2}{\|\mathbf{x}\|_2}$ را محاسبه نمایید.

ج) ماتریس $\mathbf{A} \in \mathbb{M}_n(\mathbb{R})$ را در نظر بگیرید. با استفاده از تجزیه به مقادیر تکین، پاسخ مساله بهینه‌سازی $\min\limits_{\alpha,\mathbf{Q}} \|\mathbf{A} - \alpha\mathbf{Q}\|$ را به‌گونه‌ای بیابید که $\alpha \in \mathbb{R}$ و $\mathbf{Q}$ ماتریسی یکانی (*unitary*) باشد.

د) برای ماتریس $\mathbf{A} \in \mathbb{M}_n(\mathbb{C})$، نشان دهید: $\lambda_{max}\left(\frac{\mathbf{A}+\mathbf{A}^H}{2}\right) \leq \sigma_{max}(\mathbf{A})$.

۶- باز هم مقادیر ویژه!

ماتریس هرمیتی $\mathbf{A} \in \mathbb{M}_n(\mathbb{C})$ را با عناصر $a_{i,j}$ در نظر بگیرید که همه عناصر قطری آن ۱ هستند. اگر عناصر $\mathbf{A}$ در رابطه زیر صدق کنند:

$$\sum_{j=1}^{n} |a_{i,j}| \leq 2, \qquad i = 1,2, \dots, n$$

آنگاه نشان دهید:

الف) $\mathbf{A}$ ماتریسی مثبت نیمه‌معین است.

ب) کلیه مقادیر ویژه $\mathbf{A}$ کوچکتر یا مساوی ۲ هستند، یعنی: $\lambda \leq 2$

ج) $0 \leq \det(\mathbf{A}) \leq 1$

۷- تقریب رتبه پایین ماتریس

یکی از کاربردهای تجزیه به مقادیر تکین، محاسبه تقریب رتبه پایین یک ماتریس است. ماتریس $\mathbf{A} \in \mathbb{R}^{m \times n}$ با رتبه r را با تجزیه به مقادیر تکین $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$ در نظر بگیرید. به ازای یک عدد $k \leq r$، ماتریس $\mathbf{A}_k \in \mathbb{R}^{m \times n}$ را به صورت $\mathbf{A}_k = \mathbf{U}_k\mathbf{\Sigma}_k\mathbf{V}_k^T$ تعریف می‌کنیم که در آن $\mathbf{U}_k$ و $\mathbf{V}_k$ به ترتیب k ستون اول $\mathbf{U}$ و $\mathbf{V}$ بوده و $\mathbf{\Sigma}_k$ بلوک $k \times k$ بالا و سمت چپ ماتریس $\mathbf{\Sigma}$ است که شامل k مقدار تکین بزرگ $\mathbf{A}$ می‌شود. به عبارت دیگر، $\mathbf{A}_k$ را می‌توان به صورت $\mathbf{A}_k = \sum_{i=1}^{k} \sigma_i \, \mathbf{u}_i \mathbf{v}_i^T$ بیان کرد که ماتریسی رتبه k است.

الف) بهترین تقریب رتبه k از $\mathbf{A}$ با معیار نُرم فروبینیوس از مساله بهینه‌سازی زیر به‌دست می‌آید:

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

نشان دهید که $\mathbf{A}_k$ پاسخ مساله بهینه‌سازی فوق است؛ یعنی به ازای هر ماتریس دلخواه $\mathbf{X}$ با رتبه k داریم: $\|\mathbf{A} - \mathbf{A}_k\|_F \leq \|\mathbf{A} - \mathbf{X}\|_F$.

ب) در صورتی‌که به جای نُرم فروبینیوس از معیار نُرم ۲ در مساله فوق استفاده کنیم، نشان دهید که $\mathbf{A}_k$ همچنان پاسخ مساله بهینه‌سازی تغییریافته خواهد بود.

۸- عدد حالت

ماتریس مربعی معکوس‌پذیر $\mathbf{A} \in \mathbb{R}^{n \times n}$ را در نظر بگیرید.

الف) بر اساس تعریف ذکرشده برای نُرم ۲ ماتریس، نشان دهید که $\|\mathbf{A}\|_2 = \sigma_{max}(\mathbf{A})$.

ب) با توجه مقادیر تکین $\mathbf{A}^{-1} = \mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{U}^T$، عبارت $\kappa_A = \|\mathbf{A}\|_2\|\mathbf{A}^{-1}\|_2$ را برحسب مقادیر تکین $\mathbf{A}$ محاسبه کرده و نشان دهید که $\kappa_A \geq 1$ است. κ_A را عدد حالت ماتریس $\mathbf{A}$ می‌نامیم و معیاری برای پایداری عددی ماتریس است.

ج) دستگاه معادلات $\mathbf{Ax} = \mathbf{b}$ را در نظر بگیرید. در صورتی‌که بردار $\mathbf{b}$ در فرآیند اندازه‌گیری دچار خطای کوچک $\mathbf{e}$ شود (به‌گونه‌ای‌که $\tilde{\mathbf{b}} = \mathbf{b} - \mathbf{e}$)، پاسخ دستگاه ($\tilde{\mathbf{x}}$) نیز دچار خطا خواهد شد. نشان دهید که خطای نسبی $\frac{\|\mathbf{x}-\tilde{\mathbf{x}}\|_2}{\|\mathbf{x}\|_2}$ بر حسب خطای نسبی اندازه‌گیری $\frac{\|\mathbf{e}\|_2}{\|\mathbf{b}\|_2}$ به صورت $\kappa_A^{-1}\frac{\|\mathbf{e}\|_2}{\|\mathbf{b}\|_2} \leq \frac{\|\mathbf{x}-\tilde{\mathbf{x}}\|_2}{\|\mathbf{x}\|_2} \leq \kappa_A\frac{\|\mathbf{e}\|_2}{\|\mathbf{b}\|_2}$ باند می‌شود. بر این اساس، اگر κ_A بزرگ باشد، خطای کوچک اندازه‌گیری می‌تواند به خطای بزرگی در پاسخ دستگاه منجر شود. در چنین شرایطی $\mathbf{A}$ را بد حالت می‌نامیم.

د) در صورتی‌که اندازه‌گیری و مدل به صورت همزمان دچار یک خطای کوچک شوند (یعنی: $\tilde{\mathbf{b}} = \mathbf{b} - \mathbf{e}$ و $\tilde{\mathbf{A}} = \mathbf{A} - \mathbf{E}$)، کران بالای فوق را بازنویسی نمایید.

ه) فرض کنید ماتریس $\mathbf{A}$ قطری‌شدنی بوده و به صورت $\mathbf{A} = \mathbf{S\Lambda S}^{-1}$ بیان شود. بردار باقی‌مانده $\mathbf{r} = \mathbf{A}\hat{\mathbf{x}} - \hat{\lambda}\hat{\mathbf{x}}$ را به‌گونه‌ای تعریف می‌کنیم که $\hat{\mathbf{x}}$ برداری ناصفر و $\hat{\lambda}$ عددی دلخواه است. نشان دهید که یک مقدار ویژه از $\mathbf{A}$ وجود دارد به‌گونه‌ای‌که $|\lambda - \hat{\lambda}| \leq \kappa_S \frac{\|\mathbf{r}\|}{\|\hat{\mathbf{x}}\|}$ باشد.

۹- ماتریس‌های جابجایی‌پذیر و قطری‌شدن همزمان

دو ماتریس $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n}$ را در نظر بگیرید به صورتی‌که رابطه $\mathbf{AB} = \mathbf{BA}$ برقرار است. نشان دهید:

الف) اگر $\mathbf{A}$، n مقدار ویژه متمایز داشته باشد، آنگاه ماتریس‌های $\mathbf{A}$، $\mathbf{B}$ و $\mathbf{AB}$ قطری‌شدنی خواهند بود.

ب) اگر $\mathbf{A}$ و $\mathbf{B}$ قطری‌شدنی باشند، ماتریس معکوس‌پذیر $\mathbf{P}$ وجود خواهد داشت به‌گونه‌ای‌که هر دو ماتریس $\mathbf{D}_1 = \mathbf{P}^{-1}\mathbf{AP}$ و $\mathbf{D}_2 = \mathbf{P}^{-1}\mathbf{B}P$ قطری باشند. در این حالت $\mathbf{A}$ و $\mathbf{B}$ را قطری‌شدنی همزمان می‌نامیم. به عبارتی می‌توانیم هر دو ماتریس را به صورت همزمان و در یک پایه مشترک قطری کنیم. آیا عکس این رابطه نیز برقرار است؟ (یعنی اگر $\mathbf{A}$ و $\mathbf{B}$ همزمان قطری‌شدنی باشند، می‌توان گفت $\mathbf{AB} = \mathbf{BA}$؟)

با استفاده از این نتیجه، می‌توان نشان داد که $e^{t(\mathbf{A}+\mathbf{B})} = e^{t\mathbf{A}}e^{t\mathbf{B}} = e^{t\mathbf{B}}e^{t\mathbf{A}}$ و لذا، ماتریس $e^{t\mathbf{A}}$ ماتریسی معکوس‌پذیر با معکوس $e^{-t\mathbf{A}}$ خواهد بود (حتی اگر $\mathbf{A}$ معکوس‌پذیر نباشد.) این گزاره صرفا برای اطلاعات بیشتر شما است و نیازی به اثبات آن نیست.

ج) برقراری شرط $\mathbf{AB} = \mathbf{BA}$ امری دشوار است و برای قطری‌سازی همزمان دو ماتریس باید به دنبال شرطی کاربردی‌تر باشیم. برای این منظور، ماتریس‌های متقارن $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n}$ را در نظر بگیرید. نشان دهید که دو ماتریس $\mathbf{A}$ و $\mathbf{B}$ به صورت هزمان قطری‌شدنی هستند، اگر ضرایب حقیقی α و β به‌گونه‌ای وجود داشته باشند که $\alpha\mathbf{A} + \beta\mathbf{B}$ مثبت معین شود. به طور خاص، نشان دهید در صورتی‌که $\mathbf{A}$ مثبت معین باشد، ماتریس معکوس‌پذیر $\mathbf{P}$ وجود خواهد داشت به‌گونه‌ای‌که $\mathbf{P}^T\mathbf{AP} = \mathbf{I}$ و $\mathbf{P}^T\mathbf{BP}$ قطری باشد.

الگوریتم‌های متنوعی برای قطری‌سازی همزمان دو ماتریس ارائه شده‌اند. در صورت علاقه‌مندی می‌توانید به مرجع زیر مراجعه کنید:

A. Yeredor, “On using exact joint diagonalization for noniterative approximate joint diagonalization,” IEEE Signal Process. Lett., vol. 12, no. 9, pp. 645–648, Sep. 2005.