به نام خداوند بخشنده مهربان

جبر خطی

دانشکده مهندسی کامپیوتر

حمیدرضا ربیعی، مریم رمضانی
بهار ۱۴۰۳

کمترین مربعات، مشتق ماتریس و بردار، فضای نرم

تمرین تئوری ششم

تاریخ انتشار: ۱۲ خرداد ۱۴۰۳

---

۱. پرسش‌های خود درمورد این تمرین را در سامانه کوئرا مطرح کنید.

۲. **سیاست ارسال با تاخیر:** شما در مجموع در طول نیم‌سال می‌توانید از ۱۶ روز تاخیر استفاده کنید. این مقدار برای تمارین تئوری و عملی به‌صورت جداگانه حساب می‌شود. تاخیرها با مقیاس ساعت محاسبه شده و به بالا گرد می‌شوند.

۳. **سیاست مشارکت دانشجویان در حل کردن تمارین:** دانشجویان می‌توانند در حل تمارین برای رفع ابهام و یا به‌دست آوردن ایده‌ی کلی با یک‌دیگر مشورت و همفکری کنند. این کار مورد تایید و تشویق تیم ارائه‌ی درس می‌باشد؛ چرا که هم‌فکری و کار گروهی می‌تواند موجب تقویت یادگیری شود. اما به‌دست آوردن جزئیات راه‌حل و نگارش پاسخ باید تماما توسط خود دانشجو انجام شود. **حتما در انتهای پاسخ‌های ارسالی خود نام افرادی که با آن‌ها همفکری کردید را ذکر کنید.**

پرسش ۱ (۲۵ نمره)

(آ) (۱۲ نمره) اگر $c \in \mathbb{R}$ و A یک ماتریس $n \times n$ باشد به طوری که مجموع درایه های هر ستون c باشد، نشان دهید c یک مقدار ویژه است.

(ب) (۱۳ نمره) فرض کنید $\lambda_1, \lambda_2, ..., \lambda_n$ اعداد حقیقی متمایزی باشند. ثابت کنید که مجموعه توابع $e^{\lambda_1 x}, e^{\lambda_2 x}, ..., e^{\lambda_n x}$ در فضای برداری توابع حقیقی روی $\mathbf{R}$ مستقل خطی‌اند. (راهنمایی: فرض کنید $V = span(e^{\lambda_1 x}, e^{\lambda_2 x}, ..., e^{\lambda_n x})$ باشد و عملگر T را روی V طوری تعریف کنید که به ازای هر تابع $f \in V$ داشته باشیم $Tf = f'$)

پاسخ

(آ) (۱۲ نمره)

دقت کنید که مقدار ویژه های A و A^T برابرند. لذا کافیست ثابت کنیم که c مقدار ویژه A^T است. بردار $v = [۱, \ldots, ۱]^T$ را در نظر بگیرید.

$$A^T v = A^T \begin{bmatrix} ۱ \\ ۱ \\ \vdots \\ ۱ \end{bmatrix} = \begin{bmatrix} c \\ c \\ \vdots \\ c \end{bmatrix} = cv$$

هر درایه از ضرب حاصل شده در واقع مجموع درایه های یک سطر A^T است که طبق فرض سوال برابر c است. در نتیجه c یک مقدار ویژه است.

(ب) (۱۳ نمره) فرض می‌کنیم $V = span(e^{\lambda_1 x}, e^{\lambda_2 x}, ..., e^{\lambda_n x})$ باشد. آنگاه عملگر T را روی V طوری تعریف می‌کنیم که به ازای هر تابع داخل V داشته باشیم $Tf = f'$. حال اگر $f = e^{\lambda_i x}$ را در رابطه بالا جایگذاری کنیم، خواهیم داشت:

$$Te^{\lambda_i x} = \lambda_i e^{\lambda_i x}$$

پس می‌توان دید که λ_i یک مقدار ویژه برای T بوده و بردار ویژه متناظر آن نیز $e^{\lambda_i x}$ است. حال از آنجایی که طبق فرض سوال $\lambda_۱, ..., \lambda_n$ اعداد حقیقی و متمایز هستند، پس بردار ویژه‌های متناظر با آن‌ها نیز ($(e^{\lambda_۱ x}, ..., e^{\lambda_n x})$) مستقل خطی‌اند.

پرسش ۲ (۲۵ نمره) مسئله بهینه‌سازی زیر را در نظر بگیرید:

$$\min_x \|\mathbf{W}_۱(\mathbf{A}x - b)\|_۲^۲ + \|\mathbf{W}_۲(x - c)\|_۲^۲$$

که در آن $\mathbf{A}$ ،$\mathbf{W}_۱$، و $\mathbf{W}_۲$ از جنس ماتریس و x، b و c از جنس بردار هستند. x بهینه را بیابید.

پاسخ

$$f(x) = (\mathbf{A}x - b)^T \mathbf{W}_۱^T \mathbf{W}_۱ (\mathbf{A}x - b) + (x - c)^T \mathbf{W}_۲^T \mathbf{W}_۲ (x - c)$$

$$f(x) = x^T \mathbf{A}^T \mathbf{W}_۱^T \mathbf{W}_۱ \mathbf{A}x - ۲b^T \mathbf{W}_۱^T \mathbf{W}_۱ \mathbf{A}x + b^T \mathbf{W}_۱^T \mathbf{W}_۱ b + x^T \mathbf{W}_۲^T \mathbf{W}_۲ x - ۲c^T \mathbf{W}_۲^T \mathbf{W}_۲ x + c^T \mathbf{W}_۲^T \mathbf{W}_۲ c$$

حالا گرادیان می‌گیریم و آن را برابر ۰ می‌گذاریم :

$$\nabla f = ۲\mathbf{A}^T \mathbf{W}_۱^T \mathbf{W}_۱ \mathbf{A}x - ۲\mathbf{A}^T \mathbf{W}_۱^T \mathbf{W}_۱ b + ۲\mathbf{W}_۲^T \mathbf{W}_۲ x - ۲\mathbf{W}_۲^T \mathbf{W}_۲ c = ۰$$

$$(\mathbf{A}^T \mathbf{W}_۱^T \mathbf{W}_۱ \mathbf{A} + \mathbf{W}_۲^T \mathbf{W}_۲)x = \mathbf{A}^T \mathbf{W}_۱^T \mathbf{W}_۱ b + \mathbf{W}_۲^T \mathbf{W}_۲ c$$

$$x = (\mathbf{A}^T \mathbf{W}_۱^T \mathbf{W}_۱ \mathbf{A} + \mathbf{W}_۲^T \mathbf{W}_۲)^{-۱}(\mathbf{A}^T \mathbf{W}_۱^T \mathbf{W}_۱ b + \mathbf{W}_۲^T \mathbf{W}_۲ c)$$

پرسش ۳ (۳۰ نمره)

(آ) (۱۰ نمره)

فرض کنید بردار $\mathbf{y}_{n\times ۱}$ و بردار $\mathbf{x}_{n\times ۱}$ هر دو تابعی بر حسب بردار $\mathbf{z}_{n\times ۱}$ باشند.

نشان دهید که اگر $\alpha = \mathbf{y}^T A \mathbf{x}$ باشد به طوری که A یک ماتریس $n \times n$ مستقل از $\mathbf{z}$ باشد خواهیم داشت :

$$\frac{\partial \alpha}{\partial \mathbf{z}} = \mathbf{x}^T A^T \frac{\partial \mathbf{y}}{\partial \mathbf{z}} + \mathbf{y}^T A \frac{\partial \mathbf{x}}{\partial \mathbf{z}}$$

(ب) (۱۵ نمره)

اگر A, B, C, X چهار ماتریس باشند به طوری که $F = tr[AXBXC^T]$ تعریف شده باشد.

مقدار عبارت $\dfrac{\partial tr[AXBXC^T]}{\partial X}$ را به دست آورید.

پاسخ

(آ) (۱۰ نمره) فرض کنید بردارهای X و Y دو بردار $n \times ۱$ باشند به طوری که هر دو تابعی بر حسب بردار Z باشند. اگر $\alpha = Y^T X$،می‌دانیم :

$$\alpha = \sum_{i=۱}^{n} x_i y_i \Longrightarrow \frac{\partial \alpha}{\partial z_k} = \sum_{i=۱}^{n}(x_i \frac{\partial y_i}{\partial z_k} + y_i \frac{\partial x_i}{\partial z_k}) \Longrightarrow \frac{\partial \alpha}{\partial Z} = X^T \frac{\partial Y}{\partial Z} + Y^T \frac{\partial X}{\partial Z}$$

تعریف میکنیم :

$$\mathbf{w}^T = \mathbf{y}^T A$$

با توجه به فرض سوال داریم:

$$\alpha = \mathbf{w}^T \mathbf{x}$$

با توجه به نتیجه به دست آمده از بخش قبل داریم:

$$\frac{\partial \alpha}{\partial \mathbf{z}} = \mathbf{x}^T \frac{\partial \mathbf{w}}{\partial \mathbf{z}} + \mathbf{w}^T \frac{\partial \mathbf{x}}{\partial \mathbf{z}}$$

با جایگذاری $\mathbf{w}$ خواهیم داشت :

$$\frac{\partial \alpha}{\partial \mathbf{z}} = \mathbf{x}^T A^T \frac{\partial \mathbf{y}}{\partial \mathbf{z}} + \mathbf{y}^T A \frac{\partial \mathbf{x}}{\partial \mathbf{z}}$$

(ب) (۱۵ نمره)
ابتدا در نظر بگیرید که $F = tr[AXB]$ باشد. آنگاه داریم :

$$F = \sum_i [AXB]_{ii} = \sum_i \sum_j A_{ij}[XB]_{ji} = \sum_i \sum_j A_{ij} \sum_k X_{jk} B_{ki} = \sum_i \sum_j \sum_k A_{ij} X_{jk} B_{ki}$$

$$\frac{\partial F}{\partial X_{jk}} = \sum_i A_{ij} B_{ki} = [BA]_{kj}$$

$$\Longrightarrow \frac{\partial tr[AXB]}{\partial X} = A^T B^T$$

حال با استفاده از نتیجه به دست آمده خواهیم داشت :

$$\frac{\partial tr[AXBXC^T]}{\partial X} = \frac{\partial tr[AXD]}{\partial X} + \frac{\partial tr[EXC^T]}{\partial X} = A^T D^T + E^T C$$

حال با جایگذاری $D = BXC^T$ ، $E = AXB$ داریم :

$$\frac{\partial tr[AXBXC^T]}{\partial X} = A^T C X^T B^T + B^T X^T A^T C$$

پرسش ۴ (۲۰ نمره) نشان دهید که اگر $v \in R^n$ ، $v \neq ۰$ و داشته باشیم $E \in R^{n \times n}$ آنگاه عبارت زیر برقرار خواهد بود.

$$\|E(I - \frac{vv^T}{v^T v})\|_F^۲ = \|E\|_F^۲ - \frac{\|Ev\|_۲^۲}{v^T v}$$

تعریف نرم فربنیوس : $\|A\|_F^۲ = trace[AA^T]$

(راهنمایی : خواص $orthogonal\ projection$ را برای $P := \frac{vv^T}{v^T v}$ بررسی کنید.)

پاسخ فرض کنید داشته باشیم $P := \frac{vv^T}{v^T v}$

حال با استفاده از تعریف نرم فربنیوس عبارت داده شده را ساده می‌کنیم:

$$\|E - EP\|_F^۲ = trace[(E - EP)(E - EP)^T] = trace(EE^T - EP^T E^T - EPE^T + EPP^T E^T)$$

با توجه به تعریف P می‌دانیم که یک $orthogonal\ projection$ است که نتیجه می‌دهد :

$$P^T = P, P^۲ = PP^T = P$$

$$\|E - EP\|_F^۲ = trace(EE^T - EPE^T - EPE^T + EPE^T) = trace(EE^T - EPE^T) = trace(EE^T - EP(EP)^T)$$

$$= trace(EE^T) - trace[EP(EP)^۲] = \|E\|_F^۲ - \|EP\|_F^۲$$

حال کافی است نشان دهیم : $\|EP\|_F^۲ = \frac{\|Ev\|_۲^۲}{v^T v}$

$$\|EP\|_F^۲ = trace[(EP)^T EP] = trace(P^T E^T EP) = \frac{trace(vv^T E^T E v v^T)}{(v^T v)^۲}$$

$$= \frac{v^T E^T EV \cdot v^T v}{(v^T v)^۲} = \frac{\|Ev\|_۲^۲}{v^T v}$$