به نام خداوند بخشنده مهربان

جبر خطی

دانشکده مهندسی کامپیوتر

حمیدرضا ربیعی، مریم رمضانی

پاییز ۱۴۰۲

مقدار و بردار ویژه، مشتق ماتریس، کمترین مربعات و فضای نرم — **تمرین چهارم** — تاریخ انتشار: ۲۶ آذر ۱۴۰۲

---

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. **همچنین هر تمرین تئوری و عملی را می‌توانید تا حداکثر ۳ روز با تاخیر تحویل دهید.** 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.**

سوالات تئوری (۱۵ + ۷۵ نمره)

تاریخ تحویل: ۸ دی ۱۴۰۲

پرسش ۱ (۱۲ نمره)

برای هر یک از ماتریس‌های زیر، مقادیر ویژه آن را محاسبه کنید و برای هر یک، مجموعه حداکثری از بردار ویژه‌های مستقل خطی مرتبط با آن را بیابید. در نهایت بگویید ماتریس قطری‌پذیر هست یا نه. در حالت قطری‌پذیری، یک ماتریس P بیابید که در رابطه $P^{-1}AP = D$ صدق کند.

(آ) (۴ نمره)

$$A = \begin{bmatrix} 2 & 0 & 1 \\ -1 & 2 & 3 \\ 1 & 0 & 2 \end{bmatrix}$$

(ب) (۴ نمره)

$$A = \begin{bmatrix} 6 & 2 & 0 \\ 2 & 3 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

(ج) (۴ نمره)

$$A = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

پاسخ

(آ) (۴ نمره) معادله مشخصه A را داریم:

$$det(A - \lambda I) = (2 - \lambda)(\lambda^2 - 4\lambda + 3) = 0$$

پس مقادیر ویژه برابر با ۱ و ۲ و ۳ هستند. هر مقدار ویژه یک بردار ویژه مربوط به خود را دارد

$$\lambda = 1 \rightarrow eigenvector = \begin{bmatrix} 1 \\ 4 \\ -1 \end{bmatrix}.$$

$$\lambda = 2 \rightarrow eigenvector = \begin{bmatrix} 0 \\ -1 \\ 01 \end{bmatrix}.$$

$$\lambda = 3 \rightarrow eigenvector = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}.$$

حال چون سه بردارویژه مستقل خطی داریم؛ آنگاه A diagonalizable است.ماتریس P با گرفتن بردار های ویژه و جایگزاری آنها به عنوان ستون های P بدست می آید.

$$P^{-۱}AP = \begin{bmatrix} ۱ & ۰ & ۱ \\ ۴ & -۱ & ۲ \\ -۱ & ۰ & ۱ \end{bmatrix}^{-۱} \begin{bmatrix} ۲ & ۰ & ۱ \\ -۱ & ۲ & ۳ \\ ۱ & ۰ & ۲ \end{bmatrix} \begin{bmatrix} ۱ & ۰ & ۱ \\ ۴ & -۱ & ۲ \\ -۱ & ۰ & ۱ \end{bmatrix} = \begin{bmatrix} ۱ & ۰ & ۰ \\ ۰ & ۲ & ۰ \\ ۰ & ۰ & ۳ \end{bmatrix} = D$$

(ب) (۴ نمره) معادله مشخصه A را داریم:

$$det(A - \lambda I) = (۱ + \lambda)(\lambda^۲ - ۹\lambda + ۱۴) = ۰$$

پس مقادیر ویژه برابر با −۱ و ۲ و ۷ هستند. هر مقدار ویژه یک بردار ویژه مربوط به خود را دارد.

$$\lambda = -۱ \rightarrow eigenvector = \begin{bmatrix} ۰ \\ ۰ \\ ۱ \end{bmatrix}.$$

$$\lambda = ۲ \rightarrow eigenvector = \begin{bmatrix} ۱ \\ -۲ \\ ۰ \end{bmatrix}.$$

$$\lambda = ۷ \rightarrow eigenvector = \begin{bmatrix} ۲ \\ ۱ \\ ۰ \end{bmatrix}.$$

حال چون سه بردارویژه مستقل خطی داریم؛ آنگاه A diagonalizable است.ماتریس P با گرفتن بردار های ویژه و جایگزاری آنها به عنوان ستون های P بدست می آید.

$$P^{-۱}AP = \begin{bmatrix} ۰ & ۱ & ۲ \\ ۰ & -۲ & ۱ \\ ۱ & ۰ & ۰ \end{bmatrix}^{-۱} \begin{bmatrix} ۶ & ۲ & ۰ \\ ۲ & ۳ & ۰ \\ ۰ & ۰ & -۱ \end{bmatrix} \begin{bmatrix} ۰ & ۱ & ۲ \\ ۰ & -۲ & ۱ \\ ۱ & ۰ & ۰ \end{bmatrix} = \begin{bmatrix} -۱ & ۰ & ۰ \\ ۰ & ۲ & ۰ \\ ۰ & ۰ & ۷ \end{bmatrix} = D$$

(ج) (۴ نمره) معادله مشخصه A را داریم:

$$det(A - \lambda I) = \lambda(۲ - \lambda)(\lambda - ۳) = ۰$$

پس مقادیر ویژه برابر با ۰ و ۲ و ۳ هستند. هر مقدار ویژه یک بردار ویژه مربوط به خود را دارد

$$\lambda = ۰ \rightarrow eigenvector = \begin{bmatrix} ۱ \\ ۱ \\ -۲ \end{bmatrix}.$$

$$\lambda = ۲ \rightarrow eigenvector = \begin{bmatrix} ۱ \\ -۱ \\ ۰ \end{bmatrix}.$$

$$\lambda = ۳ \rightarrow eigenvector = \begin{bmatrix} ۱ \\ ۱ \\ ۱ \end{bmatrix}.$$

حال چون سه بردارویژه مستقل خطی داریم؛ آنگاه A diagonalizable است.ماتریس P با گرفتن بردار های ویژه و جایگزاری آنها به عنوان ستون های P بدست می آید.

$$P^{-۱}AP = \begin{bmatrix} ۱ & ۱ & ۱ \\ ۱ & -۱ & ۱ \\ -۲ & ۰ & ۱ \end{bmatrix}^{-۱} \begin{bmatrix} ۲ & ۰ & ۱ \\ ۰ & ۲ & ۱ \\ ۱ & ۱ & ۱ \end{bmatrix} \begin{bmatrix} ۱ & ۱ & ۱ \\ ۱ & -۱ & ۱ \\ -۲ & ۰ & ۱ \end{bmatrix} = \begin{bmatrix} ۰ & ۰ & ۰ \\ ۰ & ۲ & ۰ \\ ۰ & ۰ & ۳ \end{bmatrix} = D$$

پرسش ۲ (۱۰ نمره) فرض کنید V یک فضای خطی روی $\mathbf{R}$ است و T تابعی خطی روی V.

(آ) (۳ نمره) فرض کنید $u, v \in V$ و u و v و $u+v$ بردارویژه‌های T هستند. نشان دهید مقدارویژه‌های متناظر u و v برابرند.

(ب) (۴ نمره) فرض کنید $Rank(T) = r$. نشان دهید T حداکثر $r+1$ مقدارویژه متمایز دارد.

(ج) (۳ نمره) فرض کنید هر عضو ناصفر V بردارویژه T است. نشان دهید T مضرب اسکالر همانی است.

پاسخ

(آ) (۳ نمره)

$$
\begin{aligned}
&T(u) = \lambda_1 u \\
&T(v) = \lambda_1 v \\
&T(u+v) = \lambda_3 (u+v) \\
&T(u+v) = T(u) + T(v) = \lambda_1 u + \lambda_2 v = \lambda_3 (u+v)
\end{aligned}
$$

برهان خلف، فرض می‌کنیم $\lambda_1 \neq \lambda_2$

$$
\begin{aligned}
&\lambda_1 = \lambda_2 + c,\ c \neq 0 \\
&\rightarrow (\lambda_2 + c)u + \lambda_2 v = \lambda_3 (u+v) \rightarrow cu = (\lambda_3 - \lambda_2)(u+v) \\
&\rightarrow T(cu) = T((\lambda_3 - \lambda_2)(u+v)) \rightarrow cT(u) = (\lambda_3 - \lambda_2) T(u+v) \rightarrow c\lambda_1 u = (\lambda_3 - \lambda_2)\lambda_3 (u+v) \\
&\rightarrow \lambda_1 (cu) = \lambda_3 ((\lambda_3 - \lambda_2)(u+v)) \rightarrow \lambda_1 = \lambda_3 \\
&\rightarrow cu = (\lambda_1 - \lambda_2)(u+v) \rightarrow cv = 0 \\
&\rightarrow c = 0 \rightarrow \lambda_1 = \lambda_2
\end{aligned}
$$

(ب) (۴ نمره) می‌دانیم که $Rank(T) = r$ پس T حداکثر می‌تواند r بردار مستقل خطی داشته باشد. از طرفی می‌دانیم که اگر x مقدار ویژه متمایز داشته باشیم بردارهای ویژه متناظرشان متمایز هستند. همچنین اگر $N(T)$ تهی نباشد می‌توانیم مقدار ویژه صفر را هم برای آن در نظر بگیریم. پس ما حداکثر r بردار مستقل خطی داریم که می‌توانند r مقدار ویژه متمایز داشته باشند که به علاوه مقدار ویژه ۰ در نهایت به $r+1$ مقدار ویژه می‌رسیم.

(ج) (۳ نمره)
طبق بخش الف می‌دانیم:

$$
u, v \in V \rightarrow T(u) = \lambda u,\ T(v) = \lambda v
$$

چون عبارت بالا برای تمام بردارهای داحل v برقرار است پس T فقط دارای یک مقدار ویژه λ است. حالا دو حالت برای λ داریم که صفر یا غیر صفر باشد.

$$
\begin{aligned}
&if\ \lambda = 0 \\
&T(x) = Ax \\
&\forall x \in V \rightarrow T = \lambda x = 0 \rightarrow A = 0 \\
&\ \rightarrow A = 0 \times I
\end{aligned}
$$

$$
\begin{aligned}
&if\ \lambda \neq 0 \\
&T(x) = Ax \\
&\forall x \in V:\ T = \lambda x = Ax \\
&\ \rightarrow\ \forall x \in V:\ (A - \lambda I)x = 0 \\
&A - \lambda I = 0 \rightarrow A = \lambda I
\end{aligned}
$$

حکم سوال اثبات شد .

پرسش ۳ (۸ نمره) p-norm را در فضای R^n به صورت $\|x\|_p = \left(\sum_{k=1}^{n} |x_k|^p\right)^{1/p}$ تعریف می‌کنیم. برای q > p > ۰ می‌توان نشان داد $\|x\|_p \geq \|x\|_q$.

(آ) (۳ نمره) نشان دهید برای یک مقدار ثابت n داریم:
$\|x\|_1 \leq \sqrt{n}\|x\|_2$

(ب) (۵ نمره) آیا می‌توان برای ثوابت p و q که q > p > ۰ است نشان داد رابطه زیر برقرار است؟ توضیح دهید.
$\|x\|_p \leq C\|x\|_q$

پاسخ

(آ) با استفاده از نامساوی کوشی شوارتز داریم:

$$||x||_1 = \sum_{i=1}^{n} |x_i|.1 \leq (\sum_{i=1}^{n} x_i^2)^{\frac{1}{2}}.(\sum_{i=1}^{n} 1^2)^{\frac{1}{2}} = \sqrt{n}||x||_2$$

(ب) با استفاده از نامساوی هولدر، داریم:

$$\sum_{i=1}^{n} |a_i||b_i| \leq (\sum_{i=1}^{n} |a_i|^r)^{\frac{1}{r}}.(\sum_{i=1}^{n} |b_i|^{\frac{r}{r-1}})^{1-\frac{1}{r}}$$

این نامساوی را برای $|a_i| = |x_i|^p$ و $|b_i| = 1$ و $r = \frac{q}{p} > 1$ اعمال می‌کنیم:

$$\sum_{i=1}^{n} |x_i|^p = \sum_{i=1}^{n} |x_i|^p.1 \leq (\sum_{i=1}^{n} (|x_i|^p)^{\frac{q}{p}})^{\frac{p}{q}}.(\sum_{i=1}^{n} 1^{\frac{q}{q-p}})^{1-\frac{p}{q}} = (\sum_{i=1}^{n} |x_i|^q)^{\frac{p}{q}}.n^{1-\frac{p}{q}}$$

درنتیجه:

$$||x||_p = (\sum_{i=1}^{n} |x_i|^p)^{\frac{1}{p}} \leq ((\sum_{i=1}^{n} |x_i|^q)^{\frac{p}{q}}.n^{1-\frac{p}{q}})^{\frac{1}{p}} = (\sum_{i=1}^{n} |x_i|^q)^{\frac{1}{q}} n^{\frac{1}{p}-\frac{1}{q}} = n^{\frac{1}{p}-\frac{1}{q}}||x||_q$$

و این یعنی مقدار C برابر است با $n^{\frac{1}{p}-\frac{1}{q}}$

پرسش ۴ (۱۵ نمره) حداقل نرم اقلیدسی

یکی از پاسخ‌های کلاسیک برای دستگاه معادلات $Ax = b$ بشرطیکه ماتریس $A \in R^{n\times m}$ و $m > n$ و این ماتریس، فول‌رنک سطری است به شکل زیر است:

$$x_{min-norm} = A^T(AA^T)^{-1}b$$

با توجه به این موضوع، به سوالات زیر پاسخ دهید:

(آ) (۵ نمره) توضیح دهید که چرا فول‌رنک سطری بودن این ماتریس لازم است؟ ادعای خود را ثابت کنید.

(ب) (۱۰ نمره) اثبات کنید که $x_{min-norm}$ در بین تمامی جواب‌های $Ax = b$ دارای کمترین نرم اقلیدسی است.

پاسخ

(آ) ماتریس A فول‌رنک است. پس می‌توان گفت که سطرهای A مستقل خطی هستند. حال، ادعا می‌کنیم که فول‌رنک سطری بودن این ماتریس، منجر به معکوس‌پذیری ماتریس $RightGram(A) = AA^T$ می‌شود.
اثبات: با برهان حلف پیش می‌رویم. فرض کنید که AA^T معکوس‌پذیر نیست. پس برای $v \in R^{n\times 1} \neq 0$ داریم:

$$AA^Tv = 0$$

از سمت چپ در v^T ضرب می‌کنیم:

$$v^TAA^Tv = 0 \longrightarrow (A^Tv)^T(A^Tv) = 0 \longrightarrow A^Tv = 0 \longrightarrow v^TA = 0$$

در نتیجه A دارای سطرهای مستقل خطی نیست. پس فرض خلف باطل است و حکم صحیح است.

(ب) برای این اثبات، در ابتدا نشان می‌دهیم که اختلاف هر بردار xی با $x_{min-norm}$ بر بردار $x_{min-norm}$ عمود است. بردار $x' \in R^{m\times 1}$ را در نظر بگیرید که $Ax' = b$:

$$(x_{min-norm} - x')^T x_{min-norm} = (x_{min-norm} - x')^T A^T(AA^T)^{-1}b =$$
$$(A(x_{min-norm} - x'))^T(AA^T)^{-1}b = (Ax_{min-norm} - Ax')^T(AA^T)^{-1}b = (b-b)^T(AA^T)^{-1}b = 0$$
$$\longrightarrow (x_{min-norm} - x') \perp x_{min-norm}$$

حالا، داریم:

$$||x'||^2 = ||x' - x_{min-norm} + x_{min-norm}||^2 = ||x' - x_{min-norm}||^2 + ||x_{min-norm}||^2 \geq ||x_{min-norm}||^2$$

اثبات کردیم که نرم اقلیدسی هر پاسخ درستی از دستگاه $Ax = b$ قطعا از نرم اقلیدسی $x_{min-nomr}$ بزرگتر یا مساوی است.

پرسش ۵ (۱۵ نمره) همانطور که می‌دانید در مسئله کمترین مربعات، به دنبال کمینه کردن جمله زیر هستیم:

$$||Ax - b||^2 = \sum_{i=1}^{m} (a_i^Tx - b_i)^2$$

که در آن a_i^T ها، سطرهای ماتریس A هستند. حالا در مسئله کمترین مربعات وزن‌دار ما به دنباله کمینه کردن جمله زیر هستیم:

$$\sum_{i=1}^{m} w_i(a_i^T x - b_i)^2$$

که w_i ها وزن‌های مثبتی هستند. این وزن‌ها به ما این قابلیت را می‌دهند که به بردارهای اختلاف $a_i^T x - b_i$ وزن‌های متفاوتی اختصاص دهیم. با این توضیحات به سوالات زیر پاسخ دهید.

(آ) (۶ نمره) نشان دهید که عبارت $\sum_{i=1}^{m} w_i(a_i^T x - b_i)^2$ را می‌توان به صورت $||D(Ax-b)||^2$ ساده کرد که در آن D یک ماتریس قطری است. اینکار باعث می‌شود که بتوانیم مسئله کمترین مربعات وزن‌دار را به شکل مسئله کمترین مربعات استاندارد درآوریم و با کمینه کردن $||Bx-d||^2$، که در آن $B = DA$ و $d = Db$ می‌باشند، مسئله را حل کنیم.

(ب) (۴ نمره) نشان دهید اگر ستون‌های ماتریس A مستقل خطی باشند، ستون‌های ماتریس B نیز مستقل خطی هستند.

(ج) (۵ نمره) می‌دانیم جواب مسئله کمترین مربعات به صورت $\hat{x} = (A^TA)^{-1}A^Tb$ می‌باشد (با فرض اینکه Aستون‌های مستقل خطی دارد). یک رابطه مشابه برای جواب مسئله کمترین مربعات وزن‌دار بدست آورید. می‌توانید در صورت نیاز در رابطه نهایی از ماتریس $W = diag(w)$ استفاده کنید.

پاسخ از آنجایی که وزن‌ها مثبت هستند می‌توانیم عبارت را به شکل زیر درآوریم.

$$\sum_{i=1}^{m} w_i(a_i^T x - b_i)^2 = \sum_{i=1}^{m} (\sqrt{w_i}(a_i^T x - b_i))^2 = ||D(Ax-b)||^2$$

که ماتریس D به صورت زیر است.

$$D = \begin{bmatrix} \sqrt{w_1} & 0 & \ddots & 0 \\ 0 & \sqrt{w_2} & \ddots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sqrt{w_m} \end{bmatrix}$$

فرض کنید ستوهای ماتریس A مستقل خطی هستند پس داریم:

$$Ax = 0 \rightarrow x = 0$$

حالا باید همین رابطه را برای ماتریس B نشان دهیم.

$$\begin{aligned} &Bx = 0 \rightarrow DAx = 0 \rightarrow Ax = 0 \rightarrow x = 0 \\ &Bx = 0 \rightarrow x = 0 \end{aligned}$$

پس ستون‌های ماتریس B هم مستقل خطی هستند. توجه کنید چون ماتریس D یک ماتریس قطری با درایه‌های مثبت است توانستیم نتیجه بگیریم $Ax = 0$ جواب مسئله کمترین مربعات وزن‌دار به صورت زیر خواهد بود:

$$\begin{aligned} (B^TB)^{-1}B^Td &= ((DA)^T(DA))^{-1}(DA)^T(Db) \\ &= (A^TD^2A)^{-1}(A^TD^2b) \\ &= (A^TWA)^{-1}(A^TWb) \end{aligned}$$

که $W = D^2 = diag(w)$

پرسش ۶ (۱۰ نمره) یک حالت تعمیم یافته از مسئله کمترین مربعات به صورت زیر است که یک تابع Affine به جمله خطا اضافه می‌شود.

$$minimize \;\; ||Ax-b||^2 + c^Tx + d$$

در معادله بالا x یک بردار n بعدی است که ما به دنبال آن هستیم. A یک ماتریس $m \times n$، b یک بردار m بعدی، c یک بردار n بعدی و d یک عدد می‌باشند. فرضیات ما در این مسئله مشابه مسئله کمترین مربعات است و فرض می‌کنیم که ستون‌های ماتریس A مستقل خطی هستند. ابتدا این مسئله را به فرم مسئله کمترین مربعات درآورید و سپس جواب آن، $\hat{x}$ را برحسب داده‌های مسئله بدست آورید.

پاسخ فرض کنید می‌خواهیم عبارت داده شده را به صورت زیر درآوریم.

$$||Ax-b||^2 + c^T x + d = ||Ax-b+f||^2 + g$$
$$\rightarrow ||Ax-b||^2 + c^T x + d = ||Ax-b||^2 + 2f^T(Ax-b) + ||f||^2 + g$$
$$\rightarrow c^T x + d = 2(A^T f)^T x - 2f^T b + ||f||^2 + g$$

فرض کنید f را به گونه‌ای انتخاب کنیم که $2A^T f = c$ پس داریم:

$$2A^T f = c$$
$$g = d + 2f^T b - ||f||^2$$

حالا مقدار زیر را برای f در نظر می‌گیریم تا رابطه مورد نظر برقرار باشد.

$$f = \frac{1}{2}((A^T A)^{-1} A^T)^T c$$
$$\rightarrow 2A^T f = A^T((A^T A)^{-1} A^T)^T c = A^T A((A^T A)^{-1})^T c = A^T A (A^T A)^{-1} c = c$$

پس توانستیم فرم سوال را به فرم مسئله کمترین مربعات درآوریم و در نتیجه جواب به صورت زیر خواهد بود.

$$x = (A^T A)^{-1} A^T (b - f)$$

پرسش ۷ (۲۰ نمره) قصد داریم گرادیان تابع L را محاسبه کنیم برای این منظور ابتدا مقادیر خواسته شده را محاسبه نمایید و سپس با استفاده از آن ها گرادیان را محاسبه کنید.(ورودی تابع بردار x با ابعاد $D_x \times 1$ است و $y \in \{0, 1\}^k$ همچنین W_1 ماتریس ضرایب با ابعاد $n \times D_x$ است.)

$$z_1 = W_1 x + b_1$$

$$a_1 = \text{LeakyReLU}(z_1, \alpha = 0/01)$$

$$z_2 = W_2 a_1 + b_2$$

$$\hat{y} = \text{Softmax}(z_2)$$

$$L = -\sum_{i=1}^{K} y_i \ln(\hat{y}_i)$$

تعریف توابع Softmax و LeakyReLU :

$$\text{Softmax}(z_1) = \frac{\exp z_1^i}{Z} \text{ where } Z = \sum_{j=1}^{K} \exp z_1^j$$

$$\text{LeakyReLU}(x, \alpha) = \begin{cases} \alpha x & x \leq 0 \\ x & x \geq 0 \end{cases}$$

مقادیر زیر را محاسبه نمایید:

(آ) (۵ نمره)

$$\frac{\partial \hat{y}_k}{\partial z_2^k}$$

(ب) (۵ نمره)

$$\frac{\partial \hat{y}_k}{\partial z_2^{i \neq k}}$$

(ج) (۵ نمره)

$$\frac{\partial L}{\partial W_1}$$

(د) (۵ نمره)

$$\frac{\partial L}{\partial b_1}$$

پاسخ

(آ) (؟ نمره) $\frac{\partial \hat{y_k}}{\partial z_2^k} = \frac{\exp(z_2^k)Z - \exp(z_2^k)^2}{Z^2} = \frac{\exp(z_2^k)}{Z} - \frac{\exp(z_2^k)^2}{Z^2} = \hat{y_k} - \hat{y_k}^2 = \delta_1$

(ب) $\frac{\partial \hat{y_k}}{\partial z_2^i} = -\frac{\exp(z_2^k)\exp(z_2^i)}{Z^2} = -\frac{\exp(z_2^k)}{Z^2}\frac{\exp(z_2^i)}{Z^2} = -\hat{y}_k\hat{y}_i = \delta_2$

(ج) $\frac{\partial L}{\partial W_1} = \frac{\partial L}{\partial \hat{y}} \times \frac{\partial \hat{y}}{\partial z_2} \times \frac{\partial z_2}{\partial a_1} \times \frac{\partial a_1}{\partial z_1} \times \frac{\partial z_1}{\partial W_1}$

$$\frac{\partial L}{\partial z_2} = H_1 = \begin{cases} -\frac{1}{\hat{y_k}}\delta_1 & i = k \\ -\frac{1}{\hat{y_i}}\delta_2 & i \neq k \end{cases}$$

$$\frac{\partial z_2}{\partial a_1} = W_2$$

$$\frac{\partial a_1}{\partial z_1} = H_2 = \begin{cases} 0.01 & z_1^i \\ 1 & \text{otherwise} \end{cases}$$

$$\frac{\partial z_1}{\partial W_1} = x^T$$

$$\frac{\partial L}{\partial W_1} = \frac{\partial L}{\partial \hat{y}} \times \frac{\partial \hat{y}}{\partial z_2} \times \frac{\partial z_2}{\partial a_1} \times \frac{\partial a_1}{\partial z_1} \times \frac{\partial z_1}{\partial W_1} = H_1 \times W_2 \times H_2 \times x^T$$

(د) $\frac{\partial L}{\partial W_1} = \frac{\partial L}{\partial \hat{y}} \times \frac{\partial \hat{y}}{\partial z_2} \times \frac{\partial z_2}{\partial a_1} \times \frac{\partial a_1}{\partial z_1} \times \frac{\partial z_1}{\partial b_1} = H_1 \times W_2 \times H_2$

سوالات عملی (۲۵ نمره)

تاریخ تحویل: ۸ دی ۱۴۰۲

پرسش ۱ (۲۵ نمره) برای سوالات عملی، به کوئرای درس مراجعه کنید.

پاسخ