تکلیف اول   زهرا توکلی 40120063   «بسمه تعالی»   جبر خطی کاربردی

سوال ۱: نرم‌ل p- نرم یک بردار: $\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}$

نرم یک بردار باید سه خاصیت positivity و homogenity و ارضا کردن traingular inequality را داشته باشد.

بنا براین:

a. بزرگ ترین درایه بردار الزاماً خاصیت positivity را ندارد اگر عددی منفی باشد. مثلاً در بردار $\begin{bmatrix}-2\\-1\end{bmatrix}$، -1 را می توانیم به عنوان نرم بردار در نظر بگیریم.

b. قدرمطلق بزرگ ترین درایه الزاماً خاصیت homogenity را ندارد. مثلاً اگر بردار $\begin{bmatrix}-2\\-1\end{bmatrix}$ را در نظر بگیریم، نرم آن برابر $|-1|=1$ است می شود. اگر $\alpha=-2$ را در بردار ضرب کنیم، نرم بردار حاصل 4 ($-2\begin{bmatrix}-2\\-1\end{bmatrix}=\begin{bmatrix}4\\2\end{bmatrix} \rightarrow$ نرم $=|4|=4$) می شود که $4 \neq |-2|1=2$.

c. تعداد درایه های بردار هم خاصیت homogenity را ندارد. مثلاً نرم بردار $\left\|\begin{bmatrix}-2\\-1\end{bmatrix}\right\|=2$ می شود. نرم بردار $\left\|\alpha\begin{bmatrix}-2\\-1\end{bmatrix}\right\|=\left\|\begin{bmatrix}-2\alpha\\-\alpha\end{bmatrix}\right\|=2$ می شود که α هایی وجود دارد که $2 \neq |\alpha|2$. به عبارتی وقتی برداری در scalar ای ضرب می شود شهوداً انتظار داریم نرم آن به همان نسبت زیاد یا کم شود ولی تعداد درایه های آن ثابت می ماند. پس اگر این تابع را به عنوان نرم در نظر بگیریم تفاوتی بین اندازه دو بردار هم بعد قائل نمی شود و دید خوبی از اندازه بردار به ما نمی دهد.

d. وارون قدرمطلق کوچک ترین درایه بردار هم خاصیت homogenity را ندارد. مثلاً بردار $\begin{bmatrix}0.001\\0.001\end{bmatrix}$ برابر 1000 می شود و نرم بردار $\begin{bmatrix}1\\1\end{bmatrix}$ برابر 1 می شود در حالی که با شهود ما از اندازه بردار مطابقت ندارد هم چنین اگر α را 1000 در نظر بگیریم $\left\|1000\begin{bmatrix}0.001\\0.001\end{bmatrix}\right\|=\left\|\begin{bmatrix}1\\1\end{bmatrix}\right\|=1 \neq 1000\times1000$ است.

e. مجموع درایه های بردار هم الزاماً خاصیت positivity را ندارد. مثلاً اگر همه درایه های بردار منفی باشند. و علاوه براین نرم بردار $\begin{bmatrix}1000\\-1000\end{bmatrix}$ از نرم $\begin{bmatrix}1\\2\end{bmatrix}$ کم تر می شود در حالی که اندازه آن شهوداً بزرگ تر است.

f. حاصل ضرب قدرمطلق درایه های بردار هم خاصیت homogenity را ندارد. مثلاً $\left\|\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}\right\|=|x_1||x_2||x_3|=|x_1x_2x_3|$ و $\left\|\begin{bmatrix}\alpha x_1\\\alpha x_2\\\alpha x_3\end{bmatrix}\right\|=|\alpha x_1\alpha x_2\alpha x_3|=|\alpha|^3|x_1x_2x_3|$ که برابر با $|\alpha||x_1x_2x_3|$ برای $\alpha \neq \pm1$ و $\alpha=0$ نیست.

سوال ۲:

3.4:

a. $(a+b)^T(a-b)=\|a\|^2-\|b\|^2$

$\rightarrow (a+b)^T(a-b)=a^Ta-\underbrace{a^Tb+b^Ta}_{0}-b^Tb=a^Ta-b^Tb=\|a\|^2-\|b\|^2$

b. $\|a+b\|^2+\|a-b\|^2=2(\|a\|^2+\|b\|^2)$

$(a+b)^T(a+b)+(a-b)^T(a-b)=a^Ta+\cancel{a^Tb}+\cancel{b^Ta}+b^Tb+a^Ta-\cancel{a^Tb}-\cancel{b^Ta}+b^Tb$

$=2a^Ta+2b^Tb=2(a^Ta+b^Tb)=2(\|a\|^2+\|b\|^2)$

۳.۱۴: اگر بزرگ ترین درایه بردار، درایه x_i باشد نزدیک ترین بردار یکه استاندارد به این بردار، بردار e_i است. اگر چند درایه بزرگ ترین مقدار را داشته باشند، e_i های نظیر تمام آنها نزدیک ترین همسایه های بردار خواهند بود. مثلاً در برداری شکل $\begin{bmatrix}3\\1\\2\end{bmatrix}$ که درایه اول بزرگ ترین است، بردار $\begin{bmatrix}1\\0\\0\end{bmatrix}$ بیش ترین شباهت را به آن دارد چون درایه اولش از بقیه بزرگ تر است.

اثبات: فرض کنیم x_i بزرگ ترین درایه بردار x باشد. می خواهیم ثابت کنیم $\forall j\neq i: \|x-e_i\|\leq\|x-e_j\|$
حالت تساوی وقتی رخ می دهد که $x_j=x_i$ و x_j هم از بزرگ ترین درایه ها باشد.

با برهان خلف فرض می کنیم $x_j<x_i$ و $\|x-e_i\|\geq\|x-e_j\|$ باشد:

$$\|x-e_i\|\geq\|x-e_j\|\Rightarrow\|x-e_i\|^2\geq\|x-e_j\|^2$$

$$\Rightarrow x_1^2+x_2^2+\dots+(x_i-1)^2+\dots+x_n^2\geq x_1^2+x_2^2+\dots(x_j-1)^2+\dots+x_n^2$$

جملات مشترک را از دو طرف حذف می کنیم.

$$\Rightarrow(x_i-1)^2+x_j^2\geq(x_j-1)^2+x_i^2$$

$$\Rightarrow x_i^2-2x_i+1+x_j^2\geq x_j^2-2x_j+1+x_i^2\Rightarrow-2x_i\geq-2x_j\Rightarrow x_i\leq x_j$$

با فرض ما تناقض دارد.

بنابراین نزدیک ترین بردار یکه استاندارد به بردار x بردار e_i خواهد بود.

۳.۱۵: می دانیم: $rms(x)^2=std(x)^2+avg(x)^2$ 3.5

$$\overset{3.5}{\Longrightarrow}std(x)^2\leq rms(x)^2\Rightarrow\sqrt{std(x)^2}\leq\sqrt{rms(x)^2}\Rightarrow\underset{+}{|std(x)|}\leq\underset{+}{|rms(x)|}\Rightarrow std(x)\leq rms(x)$$

$$\overset{3.5}{\Longrightarrow}avg(x)^2\leq rms(x)^2\Rightarrow\sqrt{avg(x)^2}\leq\sqrt{rms(x)^2}\Rightarrow|avg(x)|\leq\overset{+}{|rms(x)|}\Rightarrow|avg(x)|\leq rms(x)$$

۳.۱۷: $a.\ z=\alpha_1x_1+\alpha_2x_2+\dots+\alpha_kx_k$

$$\Rightarrow z=\begin{bmatrix}\alpha_1x_{11}+\alpha_2x_{21}+\dots+\alpha_kx_{k1}\\\alpha_1x_{12}+\alpha_2x_{22}+\dots+\alpha_kx_{k2}\\\vdots\\\alpha_1x_{1n}+\alpha_2x_{2n}+\dots+\alpha_kx_{kn}\end{bmatrix}\Rightarrow avg(z)=\frac{(\alpha_1x_{11}+\alpha_2x_{21}+\dots+\alpha_kx_{k1})+\dots+(\alpha_1x_{1n}+\alpha_2x_{2n}+\dots+\alpha_kx_{kn})}{n}$$

$$=\frac{\alpha_1(x_{11}+x_{12}+\dots+x_{1n})+\dots+\alpha_k(x_{k1}+x_{k2}+\dots+x_{kn})}{n}=\alpha_1\frac{(x_{11}+\dots+x_{1n})}{n}+\dots+\alpha_k\frac{(x_{k1}+\dots+x_{kn})}{n}$$

$$=\alpha_1avg(x_1)+\alpha_2avg(x_2)+\dots+\alpha_kavg(x_k)$$

$$b.\ std^2(z)=\frac{(\alpha_1x_{11}+\alpha_2x_{21}+\dots+\alpha_kx_{k1}-\alpha_1avg(x_1)-\dots-\alpha_kavg(x_k))^2+\dots+(\alpha_1x_{1n}+\dots+\alpha_kx_{kn}-\alpha_1avg(x_1)-\dots-\alpha_kavg(x_k))^2}{n}$$

$$=\frac{(\alpha_1(x_{11}-avg(x_1))+\dots+\alpha_k(x_{k1}-avg(x_k)))^2+\dots+(\alpha_1(x_{1n}-avg(x_1))+\dots+\alpha_k(x_{kn}-avg(x_k)))^2}{n}$$

اگر پرانتزها را به توان 2 برسانیم، به ازای هر بردار x_i جملاتی به صورت $\alpha_i^2(x_{i1}-avg(x_i))^2\dots\alpha_i^2(x_{in}-avg(x_i))^2$ و به ازای هر دو بردار نامتناظر جملاتی به صورت $\alpha_i\alpha_j(x_{i1}-avg(x_i))(x_{j1}-avg(x_j))\dots\alpha_i\alpha_j(x_{in}-avg(x_i))(x_{jn}-avg(x_j))$ داریم. جملات نوع دوم را اگر از $\alpha_i\alpha_j$ فاکتور بگیریم می توانیم به صورت $\alpha_i\alpha_j\tilde{x}_i^T\tilde{x}_j$ بنویسیم که چون x_i و x_j ها uncorrelated اند، این جملات برابر صفر می شوند. جملات نوع اول هم با فاکتورگیری از α_i^2 و تقسیم بر مخرج (n) برابر با $\alpha_i^2std(x_i)^2$ می شوند.

پس: $std(z)^2=\alpha_1^2std(x_1)^2+\dots+\alpha_k^2std(x_k)^2\Rightarrow std(z)=\sqrt{\alpha_1^2std(x_1)^2+\dots+\alpha_k^2std(x_k)^2}$

P4PCO
Scanned with CamScanner

: 3.24

a. فرض کنیم $x=\begin{bmatrix}1\\2\end{bmatrix}$ و $y=\begin{bmatrix}1000\\2000\end{bmatrix}$ و $z=\begin{bmatrix}3\\4\end{bmatrix}$ باشد. angle nearest neighbor بردار x بردار y است که در همان جهت قرار دارد و زاویه بین آنها 0 است ولی distance nearest neighbor آن بردار z است زیرا:

$$\|y-x\|=\left\|\begin{bmatrix}999\\1998\end{bmatrix}\right\|=\sqrt{3992004+998001}=2233.8$$

$$\|z-x\|=\left\|\begin{bmatrix}2\\2\end{bmatrix}\right\|=\sqrt{4+4}=\sqrt{8}=2.8$$

b. می‌خواهیم ثابت کنیم اگر $\|z_i\|=1$ باشد، angle nearest neighbor, distance nearest neighbor یکی می‌شوند:

فرض کنیم z_i distance nearest neighbor بردار x باشد:

$$\|x-z_i\|^2=(x-z_i)^T(x-z_i)=x^Tx-x^Tz_i-z_i^Tx+z_i^Tz_i=\|x\|^2-2x^Tz_i+\|z_i\|^2 \Rightarrow$$

چون $\|z_i\|$ و $\|x\|$ برای همهٔ z_iها یکسان است، پس وقتی یکی از این z_iها کم‌ترین فاصله با x را خواهد داشت که $-2x^Tz_i$ کمترین مقدار را داشته باشد. بنابراین برای آن z_i، $\frac{x^Tz_i}{\|x\|\|z_i\|}$ هم بیش‌ترین مقدار را نسبت به $\frac{x^Tz_j}{\|x\|\|z_j\|}$ها خواهد داشت و $\theta=\cos^{-1}\frac{x^Tz_i}{\|x\|\|z_i\|}$ کمترین θ خواهد بود. (با توجه به نزولی بودن تابع $\cos^{-1}$)

Scanned with CamScanner