FDU 高等线性代数 Homework 03

Due: Sept. 23, 2024
姓名: 雍崔扬
学号: 21307140051

Problem 1

设 m,n 为正整数，对于任意 $A \in \mathbb{C}^{m\times n}$，证明 l_∞ 范数诱导的范数 $\|A\|_\infty$ 的解析表达式是:

$$\|A\|_\infty := \max_{1\le i\le m}\sum_{j=1}^n |a_{ij}|$$

Solution:

任意给定矩阵 $A = \begin{bmatrix} a_1^{\mathrm{T}} \\ \vdots \\ a_m^{\mathrm{T}} \end{bmatrix} \in \mathbb{C}^{m\times n}$
对于任意满足 $\|x\|_\infty = 1$ 的 $x \in \mathbb{C}^n$，我们都有:

$$\begin{aligned}\|Ax\|_\infty &= \left\|\begin{bmatrix} a_1^{\mathrm{T}}x \\ \vdots \\ a_m^{\mathrm{T}}x \end{bmatrix}\right\|_\infty \\ &= \max_{1\le i\le m}|a_i^{\mathrm{T}}x| \\ &= \max_{1\le i\le m}\left|\sum_{j=1}^n a_{ij}x_j\right| \\ &\le \max_{1\le i\le m}\left\{\sum_{j=1}^n |a_{ij}||x_j|\right\} \\ &\le \|x\|_\infty \cdot \max_{1\le i\le m}\sum_{j=1}^n |a_{ij}| \\ &= 1\cdot \max_{1\le i\le m}\sum_{j=1}^n |a_{ij}| \\ &= \max_{1\le i\le m}\sum_{j=1}^n |a_{ij}|\end{aligned} \tag{1-1}$$

接下来我们取一个特殊的 x_0，来说明这个上界是可以取到的:

- 若 A 是全零矩阵，则没什么要证明的;
  因此我们可以假设 $A \neq 0_{m\times n}$
  设 $a_{k_0}^{\mathrm{T}}$ 是 $a_1^{\mathrm{T}}, a_2^{\mathrm{T}}, \ldots, a_m^{\mathrm{T}}$ 中绝对行和最大的行向量.
  取 $x_0 = [x_i^{(0)}] \in \mathbb{C}^n$，它满足:

$$x_i^{(0)} := \begin{cases} \frac{\bar{a}_{k_0,i}}{|a_{k_0,i}|} & \text{if } a_{k_0,i} \neq 0 \\ 0 & \text{if } a_{k_0,i} = 0 \end{cases}$$

$$\|x_0\|_\infty = 1$$

$$\|Ax_0\|_\infty = \max_{1\le i\le m}\left|\sum_{j=1}^n a_{ij}x_j^{(0)}\right| \ge \left|\sum_{j=1}^n a_{k_0,j}x_j^{(0)}\right| = \left|\sum_{j=1}^n |a_{k_0,j}|\right| = \sum_{j=1}^n |a_{k_0,j}| = \max_{1\le i\le m}\sum_{j=1}^n |a_{ij}|$$

因此式 (1-1) 中 $\|Ax\|_\infty \le \max\limits_{1\le i\le m} \sum\limits_{j=1}^{n} |a_{ij}|$ $(\forall\ x \in \mathbb{C}^n \text{ such that } \|x\|_\infty = 1)$ 的不等号是可以取等的.

于是我们有:

$$\|A\|_\infty = \max_{\|x\|_\infty=1} \|Ax\|_\infty = \max_{1\le i\le m} \sum_{j=1}^{n} |a_{ij}|$$

这样我们就得到了矩阵范数 $\|\cdot\|_\infty$，它称为**最大行和矩阵范数** (maximum row sum matrix norm)

Problem 2

设 n 是正整数，$\|\cdot\|$ 为 $\mathbb{C}^{n\times n}$ 上的相容范数，$S \in \mathbb{C}^{n\times n}$ 为非奇异矩阵 (我觉得相似矩阵用 S 作为记号更好).
证明由 $\|X\|_S := \|S^{-1}XS\|$ $(\forall\ X \in \mathbb{C}^{n\times n})$ 定义的函数也是 $\mathbb{C}^{n\times n}$ 上的相容范数，
而且它是由 $\mathbb{C}^n$ 上的范数 $\|x\|_S := \|Sx\|$ $(\forall\ x \in \mathbb{C}^n)$ 所诱导的.

Solution:
$\|\cdot\|_P$ 的非负性、正定性、齐次性和次可加性都可根据 $\|\cdot\|$ 的非负性、正定性、齐次性和次可加性简单得到.
不过还是证一下吧，以免被扣分 ( )

- ① 非负性: $\|A\|_S = \|S^{-1}AS\| \ge 0$
- ② 正定性: $\|A\|_S = \|S^{-1}AS\| = 0 \ \Leftrightarrow\ S^{-1}AS = 0_{n\times n} \ \Leftrightarrow\ A = 0_{n\times n}$
- ③ 齐次性: $\|\alpha A\|_S = \|S^{-1}(\alpha A)S\| = |\alpha|\|S^{-1}AS\| = |\alpha|\|A\|_S$
- ④ 次可加性:
  $\|A+B\|_S = \|S^{-1}(A+B)S\| = \|S^{-1}AS + S^{-1}BS\| \le \|S^{-1}AS\| + \|S^{-1}BS\| = \|A\|_S + \|B\|_S$

下证次可积性:

$$\begin{aligned} \|AB\|_S &= \|SABS^{-1}\| \\ &= \|(SAS^{-1})(SBS^{-1})\| \\ &\le \|SAS^{-1}\|\|SBS^{-1}\| \\ &= \|A\|_S\|B\|_S \end{aligned} \quad (\forall\ A, B \in \mathbb{C}^{n\times n})$$

下证最后一个结论:

$$\begin{aligned} \max_{\|x\|_S=1} \|Ax\|_S &= \max_{\|Sx\|=1} \|SAx\| \quad (\text{let } x = S^{-1}y) \\ &= \max_{\|y\|=1} \|SAS^{-1}y\| \\ &= \|SAS^{-1}\| \\ &= \|A\|_S \end{aligned}$$

表明由 $\mathbb{C}^n$ 上的范数 $\|x\|_S := \|Sx\|$ $(\forall\ x \in \mathbb{C}^n)$ 所诱导的矩阵范数就是
$\|A\|_S := \|SAS^{-1}\|$ $(\forall\ A \in \mathbb{C}^{n\times n})$

Problem 3

设 n 是正整数，用 $\|\cdot\|$ 表示 $\mathbb{C}^n$ 上的某个相容范数及其诱导的矩阵范数.
试证明对于任意非奇异矩阵 $A \in \mathbb{C}^{n\times n}$ 都有:

$$\|A^{-1}\|^{-1} = \min_{\|x\|=1} \|Ax\|$$

Solution:

$$
\begin{aligned}
\|A^{-1}\|^{-1} &= \left\{\max_{x\neq 0_n}\frac{\|A^{-1}x\|}{\|x\|}\right\}^{-1} \\
&= \min_{x\neq 0_n}\frac{\|x\|}{\|A^{-1}x\|} \quad (\text{let } x = Ay) \\
&= \min_{Ay\neq 0_n}\frac{\|Ay\|}{\|A^{-1}Ay\|} \\
&= \min_{y\neq 0_n}\frac{\|Ay\|}{\|y\|} \\
&= \min_{\|y\|=1}\|Ay\|
\end{aligned}
$$

Problem 4

设 $n\geq 1$ 是正整数，在 $\mathbb{C}^{n\times n}$ 上定义函数 $\|A\|_{\max} := \|\text{vec}(A)\|_\infty$
试证明 $\|\cdot\|_{\max}$ 是 $\mathbb{C}^{n\times n}$ 上的范数，但不具有相容性，并求最小的实数 α 使得 $N(\cdot) := \alpha\|\cdot\|_{\max}$ 是 $\mathbb{C}^{n\times n}$ 上的相容范数.

Solution:
由于 $\mathbb{C}^{n\times n}$ 和 $\mathbb{C}^{n^2}$ 是同构的，故 $\mathbb{C}^{n^2}$ 上的范数 $\|\cdot\|_\infty$ 应用到 $\mathbb{C}^{n\times n}$ 上得到的函数 $\|\cdot\|_{\max}$ 也是一个范数.
(或者可由 $\|\cdot\|_\infty$ 的非负性、正定性、齐次性和次可加性证明 $\|\cdot\|_{\max}$ 也满足非负性、正定性、齐次性和次可加性，从略)

可以举例说明它不满足次可乘性，因而不是相容范数:

$$
A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad A^2 = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}
$$
$$
\|A^2\|_{\max} = 2 > 1 = \|A\|_{\max}^2
$$

对于任意正实数 α，函数 $N(\cdot) := \alpha\|\cdot\|_{\max}$ 自然也是一个范数 (因为范数的正纯量倍数也是范数)
我们有:

$$
\begin{aligned}
N(AB) &= \alpha\max_{1\leq i,j\leq n}\left|\sum_{k=1}^n a_{ik}b_{kj}\right| \\
&\leq \alpha\max_{1\leq i,j\leq n}\sum_{k=1}^n |a_{ik}||b_{kj}| \\
&\leq \alpha\max_{i\leq i,j\leq n}\{n\|A\|_{\max}\|B\|_{\max}\} \\
&= \frac{\alpha}{n}\cdot n\|A\|_{\max}\cdot n\|B\|_{\max} \\
&= \frac{\alpha}{n}N(A)N(B)
\end{aligned}
\quad (A, B\in\mathbb{C}^{n\times n})
$$

注意到上述不等式都是紧的，至少在 $A = B = 1_n1_n^{\mathrm{T}}$ 时同时取等，因此上界 $\frac{\alpha}{n}N(A)N(B)$ 是紧的.
要使得 $N(\cdot)$ 满足次可乘性，就必须使得 $\frac{\alpha}{n}N(A)N(B)\geq N(A)N(B)\ \ (\forall\, A, B\in\mathbb{C}^{n\times n})$
因此 $\alpha\geq n$，表明所求的最小的实数 $\alpha = n$

Problem 5

设 m, n, q 是正整数，试证明:

$$
\|AB\|_{\mathrm{F}}\leq\|A\|_2\|B\|_{\mathrm{F}}\leq\|A\|_{\mathrm{F}}\|B\|_{\mathrm{F}} \ \ (\forall\, A\in\mathbb{C}^{m\times q}, B\in\mathbb{C}^{q\times n})
$$

Solution:

- 要证明 $\|A\|_2\|B\|_{\mathrm{F}} \leq \|A\|_{\mathrm{F}}\|B\|_{\mathrm{F}}\ \ (\forall\ A \in \mathbb{C}^{m\times q}, B \in \mathbb{C}^{q\times n})$
即要证明 $\|A\|_2 \leq \|A\|_{\mathrm{F}}\ \ (\forall\ A \in \mathbb{C}^{m\times q})$
而这可根据 $\|A\|_2 = \sqrt{\rho(A^{\mathrm{H}}A)} \leq \sqrt{\mathrm{tr}\,(A^{\mathrm{H}}A)} = \|A\|_{\mathrm{F}}$ 得到.
- 下面证明 $\|AB\|_{\mathrm{F}} \leq \|A\|_2\|B\|_{\mathrm{F}}\ \ (\forall\ A \in \mathbb{C}^{m\times q}, B \in \mathbb{C}^{q\times n})$

$$\begin{aligned}\|AB\|_{\mathrm{F}}^2 &= \mathrm{tr}\{(AB)^{\mathrm{H}}(AB)\} \\ &= \mathrm{tr}\left(B^{\mathrm{H}}A^{\mathrm{H}}AB\right) \\ &= \mathrm{tr}\left(A^{\mathrm{H}}ABB^{\mathrm{H}}\right)\end{aligned}$$

由于 $A^{\mathrm{H}}A$ 和 BB^{H} 都是 q 阶正规矩阵，故它们可以酉对角化，记其谱分解为 $\begin{cases}A^{\mathrm{H}}A = U^{\mathrm{H}}\Lambda_a U \\ BB^{\mathrm{H}} = V^{\mathrm{H}}\Lambda_b V\end{cases}$
其中 $U, V \in \mathbb{C}^{q\times q}$ 为酉矩阵，Λ_a, Λ_b 分别是 $A^{\mathrm{H}}A$ 和 BB^{H} 特征值构成的对角阵.
将谱分解代入 $\|AB\|_{\mathrm{F}}^2$ 的表达式，我们有:

$$\begin{aligned}\|AB\|_{\mathrm{F}}^2 &= \mathrm{tr}\left(A^{\mathrm{H}}ABB^{\mathrm{H}}\right) \\ &= \mathrm{tr}\left(U^{\mathrm{H}}\Lambda_a UV^{\mathrm{H}}\Lambda_b V\right) \\ &= \mathrm{tr}\left(\Lambda_a UV^{\mathrm{H}}\Lambda_b VU^{\mathrm{H}}\right) \\ &\leq \mathrm{tr}\left(\rho(A^{\mathrm{H}}A)I_n \cdot UV^{\mathrm{H}}\Lambda_b VU^{\mathrm{H}}\right) \\ &= \rho(A^{\mathrm{H}}A)\cdot \mathrm{tr}\left(UV^{\mathrm{H}}\Lambda_b VU^{\mathrm{H}}\right) \\ &= \rho(A^{\mathrm{H}}A)\cdot \mathrm{tr}\left(\Lambda_b VU^{\mathrm{H}}UV^{\mathrm{H}}\right) \quad (\text{note that } U^{\mathrm{H}}U = I_n \text{ and } VV^{\mathrm{H}} = I_n) \\ &= \rho(A^{\mathrm{H}}A)\cdot \mathrm{tr}\left(\Lambda_b\right) \\ &= \rho(A^{\mathrm{H}}A)\cdot \mathrm{tr}\left(BB^{\mathrm{H}}\right) \\ &= \rho(A^{\mathrm{H}}A)\cdot \mathrm{tr}\left(B^{\mathrm{H}}B\right) \\ &= \|A\|_2^2\|B\|_{\mathrm{F}}^2\end{aligned}$$

因此我们有 $\|AB\|_{\mathrm{F}} \leq \|A\|_2\|B\|_{\mathrm{F}}\ \ (\forall\ A \in \mathbb{C}^{m\times q}, B \in \mathbb{C}^{q\times n})$

综上所述，我们有 $\|AB\|_{\mathrm{F}} \leq \|A\|_2\|B\|_{\mathrm{F}} \leq \|A\|_{\mathrm{F}}\|B\|_{\mathrm{F}}\ \ (\forall\ A \in \mathbb{C}^{m\times q}, B \in \mathbb{C}^{q\times n})$
事实上我们有:

$$\begin{aligned}\|AB\|_{\mathrm{F}} &\leq \min\{\|A\|_2\|B\|_{\mathrm{F}}, \|A\|_{\mathrm{F}}\|B\|_2\} \\ &\leq \max\{\|A\|_2\|B\|_{\mathrm{F}}, \|A\|_{\mathrm{F}}\|B\|_2\} \quad (\forall\ A \in \mathbb{C}^{m\times q}, B \in \mathbb{C}^{q\times n}) \\ &\leq \|A\|_{\mathrm{F}}\|B\|_{\mathrm{F}}\end{aligned}$$

Problem 6 (optional)

设 n 是正整数，给定向量 $v \in \mathbb{C}^n$ 以及实数 $p \geq 1$
在约束条件 $\|x\|_p = 1$ 下求 $f(x) = x^{\mathrm{H}}v + v^{\mathrm{H}}x$ 的最大值.

Solution:
给定 $\mathbb{C}^n$ 上的一个范数 $\|\cdot\|$，我们定义 $\mathbb{C}^n$ 上的函数 $\|\cdot\|^D$ 为 $\|y\|^D := \max\limits_{\|x\|=1}|y^{\mathrm{H}}x| = \max\limits_{x\neq 0_n}\frac{|y^{\mathrm{H}}x|}{\|x\|}$
首先证明 $\|\cdot\|^D$ 是一个范数:

- $\|\cdot\|^D$ 的非负性、正定性和齐次性都是显然的.
- 下面证明 $\|\cdot\|^D$ 满足三角不等式:

$$\begin{aligned}\|y+z\|^D &= \max_{\|x\|=1}|(y+z)^{\mathrm{H}}x| \\ &\leq \max_{\|x\|=1}(|y^{\mathrm{H}}x| + |z^{\mathrm{H}}x|) \\ &\leq \max_{\|x\|=1}|y^{\mathrm{H}}x| + \max_{\|x\|=1}|z^{\mathrm{H}}x| \\ &= \|y\|^D + \|z\|^D\end{aligned} \quad (y, z \in \mathbb{C}^n)$$

因此 $\|\cdot\|^D$ 是一个范数，我们称其为 $\|\cdot\|$ 的**对偶范数** (dual norm)

下面证明 l_p 范数 $\|\cdot\|_p$ 的对偶范数即为 $\|\cdot\|_q$ (其中 q 是 p 的共轭子标，即满足 $\frac{1}{p}+\frac{1}{q}=1$)

> **Hölder 不等式:** $|x^{\mathrm{H}}y| \leq \|x\|_p\|y\|_q \;\; (\forall\, x,y\in\mathbb{C}^n)$
> 当且仅当 $|x|^p, |y|^q$ 线性相关时取等.
> 其中 $p,q>1$ 为共轭子标，满足 $\frac{1}{p}+\frac{1}{q}=1$

对于任意 $p>1$，取共轭子标 $q>1$ (即满足 $\frac{1}{p}+\frac{1}{q}=1$)
根据 Hölder 不等式我们有:

$$|x^{\mathrm{H}}y| \leq \|x\|_p\|y\|_q$$

取等条件是 $|x|^p, |y|^q$ 线性相关:

- 当 $y=0_n$ 时，不等号对任意 $x\in\mathbb{C}^n$ 都取等
- 当 $y\neq 0_n$ 时，不等号对由 $x_i := \begin{cases} 0 & \text{if } y_i=0 \\ \frac{|y_i|^q}{\bar{y}_i\|y\|_q^{q-1}} & \text{if } y_i\neq 0 \end{cases}$ 定义的 $x=[x_i]$ 取等

于是我们有:

$$\|y\|_p^D = \max_{\|x\|_p=1}|y^{\mathrm{H}}x| = \max_{\|x\|_p=1}\|x\|_p\|y\|_q = \|y\|_q$$

从而 $\|\cdot\|_p^D = \|\cdot\|_q$
特殊地，l_2 范数的对偶范数就是它自身，即 $\|\cdot\|_2^D = \|\cdot\|_2$
事实上，l_2 范数是 $\mathbb{C}^n$ 上仅有的自对偶范数，这并非是偶然的.

现在来考虑最大化问题:

$$\begin{aligned}
\max_{\|x\|_p=1}\{x^{\mathrm{H}}v+v^{\mathrm{H}}x\} &= \max_{\|x\|_p=1}\{x^{\mathrm{H}}v+\overline{x^{\mathrm{H}}v}\} \\
&= 2\max_{\|x\|_p=1}\mathrm{Re}(x^{\mathrm{H}}v) \\
&= 2\max_{|\alpha|=1}\max_{\|\frac{x}{\alpha}\|_p=1}\mathrm{Re}(x^{\mathrm{H}}v) \\
&= 2\max_{|\alpha|=1}\max_{\|x\|_p=1}\mathrm{Re}((\alpha x)^{\mathrm{H}}v) \\
&= 2\max_{\|x\|_p=1}\max_{|\alpha|=1}\mathrm{Re}((\alpha x)^{\mathrm{H}}v) \\
&= 2\max_{\|x\|_p=1}|x^{\mathrm{H}}v| \\
&= 2\|v\|_p^D \\
&= 2\|v\|_q
\end{aligned}$$

其中 q 是 p 的共轭子标，即满足 $\frac{1}{p}+\frac{1}{q}=1$

Problem 7 (optional)

(Hölder 不等式的推广)
设 m,n 为正整数，$A\in\mathbb{C}^{m\times n}$，且 $p_1,\ldots,p_m\in[1,+\infty]$ 满足 $\frac{1}{p_1}+\cdots+\frac{1}{p_m}=1$
试证明:

$$\left|\sum_{j=1}^n\prod_{i=1}^m a_{ij}\right| \leq \prod_{i=1}^m\left(\sum_{j=1}^n|a_{ij}|^{p_i}\right)^{\frac{1}{p_i}}$$

Solution:

对凸函数 $\begin{cases} f(x) = -\log(x) \\ \mathrm{dom}(f) = \mathbb{R}_{++} \end{cases}$ 应用 Jensen 不等式可得:

$$-\log\left(\sum_{i=1}^{m} \frac{1}{p_i} x_i\right) \leq -\sum_{i=1}^{m} \frac{1}{p_i} \log(x_i) = -\log\left(\prod_{i=1}^{m} x_i^{\frac{1}{p_i}}\right)$$

不等式两边取指数，则有:

$$\sum_{i=1}^{m} \frac{1}{p_i} x_i \geq \prod_{i=1}^{m} x_i^{\frac{1}{p_i}}$$

给定 $j \in \{1, \ldots, n\}$，令 $x_i = \frac{|a_{ij}|^{p_i}}{\sum_{j=1}^{n} |a_{ij}|^{p_i}}$，则我们有:

$$\sum_{i=1}^{m} \left\{ \frac{1}{p_i} \frac{|a_{ij}|^{p_i}}{\sum_{j=1}^{n} |a_{ij}|^{p_i}} \right\} \geq \prod_{i=1}^{m} \left\{ \left(\frac{|a_{ij}|^{p_i}}{\sum_{j=1}^{n} |a_{ij}|^{p_i}} \right)^{\frac{1}{p_i}} \right\}$$

左右两式对 $j = 1, \ldots, n$ 求和，则我们有:

$$\begin{aligned}
\mathrm{LHS} &= \sum_{j=1}^{n} \sum_{i=1}^{m} \left\{ \frac{1}{p_i} \frac{|a_{ij}|^{p_i}}{\sum_{j=1}^{n} |a_{ij}|^{p_i}} \right\} \\
&= \sum_{i=1}^{m} \sum_{j=1}^{n} \left\{ \frac{1}{p_i} \frac{|a_{ij}|^{p_i}}{\sum_{j=1}^{n} |a_{ij}|^{p_i}} \right\} \\
&= \sum_{i=1}^{m} \left\{ \frac{1}{p_i} \frac{\sum_{j=1}^{n} |a_{ij}|^{p_i}}{\sum_{j=1}^{n} |a_{ij}|^{p_i}} \right\} \\
&= \sum_{i=1}^{m} \left\{ \frac{1}{p_i} \cdot 1 \right\} \\
&= 1 \\
\hline
\mathrm{RHS} &= \sum_{j=1}^{n} \prod_{i=1}^{m} \left\{ \left(\frac{|a_{ij}|^{p_i}}{\sum_{j=1}^{n} |a_{ij}|^{p_i}} \right)^{\frac{1}{p_i}} \right\} \\
&= \sum_{j=1}^{n} \prod_{i=1}^{m} \left\{ \frac{|a_{ij}|}{(\sum_{j=1}^{n} |a_{ij}|^{p_i})^{\frac{1}{p_i}}} \right\} \\
&= \frac{\sum_{j=1}^{n} \prod_{i=1}^{m} |a_{ij}|}{\prod_{i=1}^{m} (\sum_{j=1}^{n} |a_{ij}|^{p_i})^{\frac{1}{p_i}}}
\end{aligned}$$

根据 $\mathrm{LHS} \geq \mathrm{RHS}$ 可知:

$$1 = \mathrm{LHS} \geq \mathrm{RHS} = \frac{\sum_{j=1}^{n} \prod_{i=1}^{m} |a_{ij}|}{\prod_{i=1}^{m} (\sum_{j=1}^{n} |a_{ij}|^{p_i})^{\frac{1}{p_i}}}$$

$$\Leftrightarrow$$

$$\sum_{j=1}^{n} \prod_{i=1}^{m} |a_{ij}| \leq \prod_{i=1}^{m} \left(\sum_{j=1}^{n} |a_{ij}|^{p_i} \right)^{\frac{1}{p_i}}$$

因此我们有:

$$
\begin{aligned}
\left|\sum_{j=1}^{n}\prod_{i=1}^{m}a_{ij}\right| &\le \sum_{j=1}^{n}\left|\prod_{i=1}^{m}a_{ij}\right| \quad (\text{triangle inequality; and note that } |ab|=|a||b| \text{ for all } a,b\in\mathbb{C})\\
&= \sum_{j=1}^{n}\prod_{i=1}^{m}|a_{ij}| \quad (\text{use the conclusion above})\\
&\le \prod_{i=1}^{m}\left(\sum_{j=1}^{n}|a_{ij}|^{p_i}\right)^{\frac{1}{p_i}}
\end{aligned}
$$

Problem 8 (optional)

(三角不等式的推广)
设 $n \geq 2$ 为正整数，$x_1,\ldots,x_n,\beta$ 都是实数.
试证明:

$$\sum_{i=1}^{n}[(x_i-\beta)^2+\beta^2]^{\frac{1}{2}} \geq (\sum_{i=1}^{n}x_i^2)^{\frac{1}{2}}$$

邵老师提供的证明:
记 $\mathbb{C}^n$ 的第 i 个标准单位基向量为 e_i，则对于任意 $x\in\mathbb{R}^n$ 和 $\beta\in\mathbb{R}$ 我们都有:

$$
\begin{aligned}
&\sum_{i=1}^{n}\left[(x_i-\beta)^2+\beta^2\right]^{\frac{1}{2}} \\
&=\|(x_1-\beta)e_1+\beta e_2\|_2+\|(x_2-\beta)e_2+\beta e_3\|_2+\cdots+\|(x_{n-1}-\beta)e_{n-1}+\beta e_n\|_2+\|(x_n-\beta)e_n+\beta e_1\|_2 \\
&=\left\|\begin{bmatrix} x_1-\beta \\ \beta \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \end{bmatrix}\right\|_2+\left\|\begin{bmatrix} 0 \\ x_2-\beta \\ \beta \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \end{bmatrix}\right\|_2+\cdots+\left\|\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ x_{n-1}-\beta \\ \beta \end{bmatrix}\right\|_2+\left\|\begin{bmatrix} \beta \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ x_n-\beta \end{bmatrix}\right\|_2 \\
&\geq\left\|\begin{bmatrix} x_1-\beta \\ \beta \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \end{bmatrix}+\begin{bmatrix} 0 \\ x_2-\beta \\ \beta \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \end{bmatrix}+\cdots+\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ x_{n-1}-\beta \\ \beta \end{bmatrix}+\begin{bmatrix} \beta \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ x_n-\beta \end{bmatrix}\right\|_2 \\
&=\left\|\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ \vdots \\ x_{n-2} \\ x_{n-1} \\ x_n \end{bmatrix}\right\|_2 \\
&=\|x\|_2 \\
&=\left(\sum_{i=1}^{n}x_i^2\right)^{\frac{1}{2}}
\end{aligned}
$$

Failed Solution:
注意到 $x=0_n$ 时不等式化为 $n\sqrt{2}|\beta|\geq 0$，显然对于任意 $\beta\in\mathbb{R}$ 都成立.
注意到 $\beta=0$ 时不等式化为 $\|x\|_1\geq\|x\|_2$，显然对任意 $x\in\mathbb{R}^n$ 都成立.

对任意给定的 $\beta\neq 0\in\mathbb{R}$，定义 $\mathbb{R}^n\mapsto\mathbb{R}$ 的函数:

$$
\begin{aligned}
f(x)&=\sum_{i=1}^{n}[(x_i-\beta)^2+\beta^2]^{\frac{1}{2}}-\left(\sum_{i=1}^{n}x_i^2\right)^{\frac{1}{2}} \\
&=\sum_{i=1}^{n}[(x_i-\beta)^2+\beta^2]^{\frac{1}{2}}-\|x\|_2 \\
&=\sum_{i=1}^{n}\left\|\begin{bmatrix} x_i-\beta \\ \beta \end{bmatrix}\right\|_2-\|x\|_2
\end{aligned}
$$

显然 $f(x)$ 是关于 $x\in\mathbb{R}^n$ 的严格凸函数，且在 $\mathbb{R}^n$ 上连续 (仅在 $x=0_n$ 处不可微)
因此其全局最小点唯一，且为驻点.

对 x 求梯度可得: (其中 e_i 代表 $\mathbb{R}^n$ 的第 i 个标准单位基向量)

$$\nabla f(x) = \sum_{i=1}^{n} [(x_i - \beta)^2 + \beta^2]^{-\frac{1}{2}} (x_i - \beta) e_i - \frac{x}{\|x\|_2}$$

令 $\nabla f(x) = 0_n$ 可知全局最小点 $x^\star$ 满足:

$$[(x_i^\star - \beta)^2 + \beta^2]^{-\frac{1}{2}} (x_i^\star - \beta) - \frac{x_i^\star}{\|x^\star\|_2} = 0 \ \ (\forall\, i = 1, \ldots, n)$$
$$\Leftrightarrow$$
$$[(x_i^\star - \beta)^2 + \beta^2]^{\frac{1}{2}} = \frac{x_i^\star - \beta}{x_i^\star} \|x^\star\|_2 \ \ (\forall\, i = 1, \ldots, n)$$
$$\Leftrightarrow$$
$$\frac{(x_i^\star)^2}{\|x^\star\|_2^2} = \frac{(x_i^\star - \beta)^2}{(x_i^\star - \beta)^2 + \beta^2} = 1 - \frac{\beta^2}{(x_i^\star - \beta)^2 + \beta^2} \ \ (\forall\, i = 1, \ldots, n)$$

因此全局最小值 $f(x^\star)$ 满足:

$$\begin{aligned} f(x^\star) &= \sum_{i=1}^{n} [(x_i^\star - \beta)^2 + \beta^2]^{\frac{1}{2}} - \|x^\star\|_2 \quad \left(\text{note that } [(x_i^\star - \beta)^2 + \beta^2]^{\frac{1}{2}} = \frac{x_i^\star - \beta}{x_i^\star} \|x^\star\|_2\right) \\ &= \sum_{i=1}^{n} \frac{x_i^\star - \beta}{x_i^\star} \|x^\star\|_2 - \|x^\star\|_2 \\ &= \|x^\star\|_2 \left\{ (n-1) - \sum_{i=1}^{n} \frac{\beta}{x_i^\star} \right\} \end{aligned}$$

要证明 $f(x^\star) \geq 0$，只需证明 $\sum_{i=1}^{n} \frac{\beta}{x_i^\star} \leq (n-1)$ 即可，但我证不出来.

The End