Math 217 Fall 2013 Homework 2 Solutions

Due Thursday Sept. 26, 2013 5pm

- This homework consists of 6 problems of 5 points each. The total is 30.
- You need to fully justify your answer – prove that your function indeed has the specified property – for each problem.
- Please read this week's lecture notes before working on the problems.

Question 1. *The following are several possible strategies to prove Cauchy-Schwarz:*

$$|\boldsymbol{x}\cdot\boldsymbol{y}| = |x_1\, y_1 + \cdots + x_N\, y_N| \leqslant (x_1^2 + \cdots + x_N^2)^{1/2}\,(y_1^2 + \cdots + y_N^2)^{1/2} = \|\boldsymbol{x}\|\,\|\boldsymbol{y}\|. \tag{1}$$

Pick any one (or come up with your own) idea and write down a detailed proof.

- *Approach 1.*
  Mathematical induction.
- *Approach 2.*
  Let $t \in \mathbb{R}$. Then $(\boldsymbol{x} - t\,\boldsymbol{y})\cdot(\boldsymbol{x} - t\,\boldsymbol{y}) \geqslant 0$ for all t. Write the left hand side as a quadratic polynomial of t.
- *Approach 3.*
  Use $x_i\, y_i = \left(\frac{x_i}{k}\right)(y_i\, k) \leqslant \frac{1}{2}\,(x_i^2\, k^{-2} + y_i^2\, k^2)$. Choose appropriate k.

Solution.

- Approach 1.

  Though the case $N = 1$ is trivial. For reasons that will be clear in a few lines, we have to prove $N = 2$. This is done in Sept. 16's lecture and is omitted here.

  Now we try to prove the case $N = k + 1$ assuming

  $$|x_1\, y_1 + \cdots + x_k\, y_k| \leqslant (x_1^2 + \cdots + x_k^2)^{1/2}\,(y_1^2 + \cdots + y_k^2)^{1/2} \tag{2}$$

  We have

  $$\begin{aligned} |x_1\, y_1 + \cdots + x_k\, y_k + x_{k+1}\, y_{k+1}| &\leqslant |x_1\, y_1 + \cdots + x_k\, y_k| + |x_{k+1}\, y_{k+1}| \\ &\leqslant (x_1^2 + \cdots + x_k^2)^{1/2}\,(y_1^2 + \cdots + y_k^2)^{1/2} + |x_{k+1}|\,|y_{k+1}| \\ &\leqslant ((x_1^2 + \cdots + x_k^2) + |x_{k+1}|^2)^{1/2}\,((y_1^2 + \cdots + y_k^2) + |y_{k+1}|^2)^{1/2} \\ &= (x_1^2 + \cdots + x_{k+1}^2)^{1/2}\,(y_1^2 + \cdots + y_{k+1}^2)^{1/2}. \end{aligned} \tag{3}$$

  Note that in the last inequality we have used the $N = 2$ case.

- Approach 2.

  Since $(\boldsymbol{x} - t\,\boldsymbol{y})\cdot(\boldsymbol{x} - t\,\boldsymbol{y}) = (\boldsymbol{y}\cdot\boldsymbol{y})\,t^2 - 2\,(\boldsymbol{x}\cdot\boldsymbol{y})\,t + (\boldsymbol{x}\cdot\boldsymbol{x})$, the fact that it is non-negative implies

  $$[2\,(\boldsymbol{x}\cdot\boldsymbol{y})]^2 - 4\,(\boldsymbol{y}\cdot\boldsymbol{y})\,(\boldsymbol{x}\cdot\boldsymbol{x}) \leqslant 0 \tag{4}$$

  which gives Cauchy-Schwarz.

- Approach 3.

  Let $k \in \mathbb{R}$ to be determined later. We have

  $$x_1\, y_1 + \cdots + x_N\, y_N \leqslant \frac{1}{2}\left[\frac{x_1^2 + \cdots + x_N^2}{k^2} + k^2\,(y_1^2 + \cdots + y_N^2)\right]. \tag{5}$$

Now take

$$k^2 = \frac{(x_1^2 + \cdots + x_N^2)^{1/2}}{(y_1^2 + \cdots + y_N^2)^{1/2}}. \tag{6}$$

The proof ends.

Question 2. *Let $E \subseteq \mathbb{R}^N$. Define its distance function $d: \mathbb{R}^N \mapsto \mathbb{R}$ as*

$$d(\boldsymbol{x}) := \inf_{\boldsymbol{y} \in E} \operatorname{dist}(\boldsymbol{x}, \boldsymbol{y}) = \inf_{\boldsymbol{y} \in E} \|\boldsymbol{x} - \boldsymbol{y}\|. \tag{7}$$

Prove that $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^N$, $|d(\boldsymbol{x}) - d(\boldsymbol{y})| \leqslant \|\boldsymbol{x} - \boldsymbol{y}\|$.

Proof. First we prove $d(\boldsymbol{x}) - d(\boldsymbol{y}) \leqslant \|\boldsymbol{x} - \boldsymbol{y}\|$. We have, for any $\boldsymbol{z} \in E$,

$$\begin{aligned} d(\boldsymbol{x}) - \operatorname{dist}(\boldsymbol{y}, \boldsymbol{z}) &= \inf_{\boldsymbol{w} \in E} \operatorname{dist}(\boldsymbol{x}, \boldsymbol{w}) - \operatorname{dist}(\boldsymbol{y}, \boldsymbol{z}) \\ &\leqslant \operatorname{dist}(\boldsymbol{x}, \boldsymbol{z}) - \operatorname{dist}(\boldsymbol{y}, \boldsymbol{z}) \\ &= \|\boldsymbol{x} - \boldsymbol{z}\| - \|\boldsymbol{y} - \boldsymbol{z}\| \\ &\leqslant \|\boldsymbol{x} - \boldsymbol{y}\|. \end{aligned} \tag{8}$$

Here we applied triangle's inequality in the last inequality. Note that $\|\boldsymbol{x} - \boldsymbol{y}\|$ is independent of $\boldsymbol{z}$. Therefore we can take infimum and obtain

$$d(\boldsymbol{x}) - d(\boldsymbol{y}) = d(\boldsymbol{x}) - \inf_{\boldsymbol{z} \in E} \operatorname{dist}(\boldsymbol{y}, \boldsymbol{z}) \leqslant \|\boldsymbol{x} - \boldsymbol{y}\|. \tag{9}$$

Finally noticing the symmetry between $\boldsymbol{x}$ and $\boldsymbol{y}$, we have

$$d(\boldsymbol{y}) - d(\boldsymbol{x}) \leqslant \|\boldsymbol{y} - \boldsymbol{x}\| = \|\boldsymbol{x} - \boldsymbol{y}\|. \tag{10}$$

Summarizaing the above, we have $|d(\boldsymbol{x}) - d(\boldsymbol{y})| \leqslant \|\boldsymbol{x} - \boldsymbol{y}\|$. ☐

Question 3.

a) *Prove that the following are both norms on $\mathbb{R}^N$:*

$$\|\boldsymbol{x}\|_\infty := \max_{i=1,\ldots,N} \{|x_i|\}; \qquad \|\boldsymbol{x}\|_1 := |x_1| + |x_2| + \cdots + |x_N|; \tag{11}$$

b) *Let X be a linear vector space with norm $\|\cdot\|$. Prove the following: If one can define an inner product $(\cdot, \cdot)$ such that $\|x\| = (x, x)^{1/2}$, then for any $x, y \in X$,*

$$\|x + y\|^2 + \|x - y\|^2 = 2\,(\|x\|^2 + \|y\|^2). \tag{12}$$

c) *Find a norm on $\mathbb{R}^N$ that cannot be defined through an inner product. Justify your answer.*

Solution.

a) We check

i. $\|\boldsymbol{x}\|_\infty := \max_{i=1,\ldots,N} \{|x_i|\} \geqslant 0$; $\|\boldsymbol{x}\|_\infty = 0 \Longrightarrow \max_i |x_i| = 0 \Longrightarrow x_i = 0$ for all $i = 1, 2, \ldots, N \Longrightarrow \boldsymbol{x} = \mathbf{0}$;

$\|\boldsymbol{x}\|_1 := |x_1| + |x_2| + \cdots + |x_N| \geqslant 0$; $\|\boldsymbol{x}\|_1 = 0 \Longrightarrow |x_1| + |x_2| + \cdots + |x_N| = 0 \Longrightarrow \|\boldsymbol{x}\|_\infty = 0 \Longrightarrow \max_i |x_i| = 0 \Longrightarrow x_i = 0$ for all $i = 1, 2, \ldots, N \Longrightarrow \boldsymbol{x} = \mathbf{0}$.

ii. $\|a\,\boldsymbol{x}\|_\infty = \max_i \{|a\,x_i|\} = \max_i \{|a|\,|x_i|\} = |a| \max_i \{|x_i|\} = |a|\,\|\boldsymbol{x}\|_\infty$;

$\|a\,\boldsymbol{x}\|_1 = |a\,x_1| + |a\,x_2| + \cdots + |a\,x_N| = |a|\,(|x_1| + |x_2| + \cdots + |x_N|) = |a|\,\|\boldsymbol{x}\|_1$.

iii. (Triangle inequality).

$$\begin{aligned} \|\boldsymbol{x}+\boldsymbol{y}\|_\infty &= \max_i |x_i+y_i| \\ &\leqslant \max_i (|x_i|+|y_i|) \\ &\leqslant \max_i |x_i| + \max_i |y_i| \\ &= \|\boldsymbol{x}\|_\infty + \|\boldsymbol{y}\|_\infty. \end{aligned} \tag{13}$$

$$\begin{aligned} \|\boldsymbol{x}+\boldsymbol{y}\|_1 &= |x_1+y_1|+|x_2+y_2|+\cdots+|x_N+y_N| \\ &\leqslant |x_1|+|y_1|+\cdots+|x_N|+|y_N| \\ &= (|x_1|+|x_2|+\cdots+|x_N|)+(|y_1|+|y_2|+\cdots+|y_N|) \\ &= \|\boldsymbol{x}\|_1+\|\boldsymbol{y}\|_1. \end{aligned} \tag{14}$$

b) We have

$$\begin{aligned} \|x+y\|^2+\|x-y\|^2 &= (x+y,x+y)+(x-y,x-y) \\ &= (x,x)+(x,y)+(y,x)+(y,y) \\ &\quad +(x,x)+(x,-y)+(-y,x)+(-y,-y) \\ &= (x,x)+2\,(x,y)+(y,y) \\ &\quad +(x,x)-2\,(x,y)+(y,y) \\ &= 2\,[(x,x)+(y,y)] \\ &= 2\,(\|x\|^2+\|y\|^2). \end{aligned} \tag{15}$$

c) Take $\|\cdot\|_\infty$. All we need to show is that it does not satisfy the equality proved in b). Take $\boldsymbol{x}=\boldsymbol{e}_1, \boldsymbol{y}=\boldsymbol{e}_2$. Then we have $\|\boldsymbol{x}+\boldsymbol{y}\|_\infty=\|\boldsymbol{x}-\boldsymbol{y}\|_\infty=\|\boldsymbol{x}\|_\infty=\|\boldsymbol{y}\|_\infty=1$. The equality is not satisfied.

Question 4. *Let $O\in\mathbb{R}^{N\times N}$ be such that $\|O\,\boldsymbol{x}\|=\|\boldsymbol{x}\|$ for any $\boldsymbol{x}\in\mathbb{R}^N$. Prove that O is orthogonal. Please prove it directly and do not use any theorem from linear algebra.*

Proof. First we show that $(O\,\boldsymbol{x})\cdot(O\,\boldsymbol{y})=\boldsymbol{x}\cdot\boldsymbol{y}$ for all $\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^N$. To see this we calculate

$$\begin{aligned} \boldsymbol{x}\cdot\boldsymbol{x}+2\,\boldsymbol{x}\cdot\boldsymbol{y}+\boldsymbol{y}\cdot\boldsymbol{y} &= (\boldsymbol{x}+\boldsymbol{y})\cdot(\boldsymbol{x}+\boldsymbol{y}) \\ &= [O(\boldsymbol{x}+\boldsymbol{y})]\cdot[O(\boldsymbol{x}+\boldsymbol{y})] \\ &= (O\,\boldsymbol{x})\cdot(O\,\boldsymbol{x})+2\,(O\,\boldsymbol{x})\cdot(O\,\boldsymbol{y})+(O\,\boldsymbol{y})\cdot(O\,\boldsymbol{y}) \\ &= \|O\,\boldsymbol{x}\|^2+2\,(O\,\boldsymbol{x})\cdot(O\,\boldsymbol{y})+\|O\,\boldsymbol{y}\|^2 \\ &= \|\boldsymbol{x}\|^2+2\,(O\,\boldsymbol{x})\cdot(O\,\boldsymbol{y})+\|\boldsymbol{y}\|^2 \\ &= \boldsymbol{x}\cdot\boldsymbol{x}+2\,(O\,\boldsymbol{x})\cdot(O\,\boldsymbol{y})+\boldsymbol{y}\cdot\boldsymbol{y}. \end{aligned} \tag{16}$$

The claim follows.

Recalling $\boldsymbol{x}\cdot\boldsymbol{y}=\boldsymbol{x}^T\boldsymbol{y}$, we have

$$(O\,\boldsymbol{x})\cdot(O\,\boldsymbol{y})=(O\,\boldsymbol{x})^T\,(O\boldsymbol{y})=\boldsymbol{x}^T O^T O\,\boldsymbol{y}=[O^T O\,\boldsymbol{x}]^T\boldsymbol{y}=(O^T O\,\boldsymbol{x})\cdot\boldsymbol{y}. \tag{17}$$

Thus we have shown

$$[(O^T O\,\boldsymbol{x})-\boldsymbol{x}]\cdot\boldsymbol{y}=0 \tag{18}$$

for all $\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^N$.

Taking $\boldsymbol{y}=\boldsymbol{e}_1,\ldots,\boldsymbol{e}_N$, we see that

$$O^T O\,\boldsymbol{x}=\boldsymbol{x} \tag{19}$$

for all $\boldsymbol{x} \in \mathbb{R}^N$.

Finally taking $\boldsymbol{x} = \boldsymbol{e}_1, \ldots, \boldsymbol{e}_N$ we see that $O^T O = I$, that is the matrix O is orthogonal. ☐

Question 5. *Let $D = \operatorname{diag}(d_1, \ldots, d_N)$ be a diagonal matrix with all the d_i's distinct. Let $A \in \mathbb{R}^{N \times N}$ be such that $A\,D = D\,A$. What can we conclude about A? Justify your answer.*

Proof. The (i, j) entry for $A\,D$ is $d_j\,a_{ij}$ while the (i, j) entry for $D\,A$ is $d_i\,a_{ij}$. Thus we have

$$(d_i - d_j)\,a_{ij} = 0 \tag{20}$$

for all $i, j = 1, \ldots, N$. As d_i's are distinct, this means $a_{ij} = 0$ when $i \neq j$, that is A is diagonal.

It is clear that if A is diagonal, then $A\,D = D\,A$. Thus we have fully characterized the matrices that commute with a diagonal matrix with distinct main diagonal entries. ☐

Question 6. (Twin Prime Conjecture) *Earlier this year, Prof. Yitang Zhang of University of New Hampshire made history through proving the following result:*

$$\liminf_{n \longrightarrow \infty} (p_{n+1} - p_n) < 7 \times 10^7 \tag{21}$$

where p_n is the n-th prime number.

a) *Prove that the Twin Prime Conjecture "There are infinitely many pairs of prime numbers with difference 2" is equivalent to*

$$\liminf_{n \longrightarrow \infty} (p_{n+1} - p_n) = 2. \tag{22}$$

b) *One step of his proof is basically the following. Assume*

$$\sum_{d < D^2, d \mid \mathcal{P}} \sum_{c \in \mathcal{C}_i(d)} |\Delta(\theta, d, c)| \leqslant x\,(\log x)^{-A}, \tag{23}$$

for some $A > 0$ and

$$\sum_{c \in \mathcal{C}_i(d)} |\Delta(\theta, d, c)| \leqslant x\,(\log x)/d; \qquad \sum_{d < D^2, d \mid \mathcal{P}} \tau_3(d)^2\,\rho_2(d)^2\,d^{-1} \leqslant (\log x)^B \tag{24}$$

for some $B > 0$. Then we have

$$\mathcal{E} := \left| \sum_{d < D^2, d \mid \mathcal{P}} \tau_3(d)\,\rho_2(d) \sum_{c \in \mathcal{C}_i(d)} |\Delta(\theta, d, c)| \right| \ll x\,(\log x)^{\frac{B+1-A}{2}}. \tag{25}$$

for any $A > 0$. Prove the above claim using Cauchy-Schwarz.

Proof.

a) If $\liminf_{n \longrightarrow \infty} (p_{n+1} - p_n) = 2$, then there is a subsequence satisfying

$$\liminf_{k \longrightarrow \infty} (p_{n_k+1} - p_{n_k}) = 2. \tag{26}$$

Consequently, there is $K \in \mathbb{N}$ such that for all $k > K$,

$$|p_{n_k+1} - p_{n_k} - 2| < 1/2. \tag{27}$$

But the left hand side is an integer, so it must be 0. That is there are infinitely many pairs of prime numbers with difference 2.

b) We have

$$\begin{aligned}
\mathcal{E} &= \left| \sum_{d<D^2, d|\mathcal{P}} \tau_3(d)\,\rho_2(d) \sum_{c\in\mathcal{C}_i(d)} |\Delta(\theta,d,c)| \right| \left(\sum_{d<D^2, d|\mathcal{P}} \sum_{c\in\mathcal{C}_i(d)} |\Delta(\theta,d,c)| \right)^{1/2} \\
&= \left| \sum_{d<D^2, d|\mathcal{P}} \sum_{c\in\mathcal{C}_i(d)} \left(\tau_3(d)\,\rho_2(d)\,|\Delta(\theta,d,c)|^{1/2}\right) \left(|\Delta\,(\theta,d,c)|^{1/2}\right) \right| \\
&\leqslant \left(\sum_{d<D^2, d|\mathcal{P}} \sum_{c\in\mathcal{C}_i(d)} \left(\tau_3(d)\,\rho_2(d)\,|\Delta(\theta,d,c)|^{1/2}\right)^2 \right)^{1/2} \left(\sum_{d<D^2, d|\mathcal{P}} \sum_{c\in\mathcal{C}_i(d)} |\Delta(\theta,d,c)| \right)^{1/2} \\
&= \left(\sum_{d<D^2, d|\mathcal{P}} \sum_{c\in\mathcal{C}_i(d)} (\tau_3(d)^2\,\rho_2(d)^2\,|\Delta(\theta,d,c)| \right)^{1/2} \left(\sum_{d<D^2, d|\mathcal{P}} \sum_{c\in\mathcal{C}_i(d)} |\Delta(\theta,d,c)| \right)^{1/2} \\
&= \left(\sum_{d<D^2, d|\mathcal{P}} \tau_3(d)^2\,\rho_2(d)^2 \left[\sum_{c\in\mathcal{C}_i(d)} |\Delta(\theta,d,c)| \right] \right)^{1/2} (x\,(\log x)^{-A})^{1/2} \\
&\leqslant \left(\sum_{d<D^2, d|\mathcal{P}} \tau_3(d)^2\,\rho_2(d)^2\,d^{-1}\,x\,(\log x) \right)^{1/2} (x\,(\log x)^{-A})^{1/2} \\
&\leqslant x^{1/2}\,(\log x)^{\frac{B+1}{2}}\,x^{1/2}\,(\log x)^{-A/2} \\
&= x\,(\log x)^{\frac{B+1-A}{2}}. \qquad (28)
\end{aligned}$$

□