Math 63CM Homework 4 Solutions

Kevin Yang

Stanford University

February 8, 2020

Problem 1

(i). First, for $x = 0$ the desired inequality is obvious.

For $x \neq 0$, we claim it suffices to show the inequality for $x \in \mathbb{R}^n$ with $\|x\| = 1$. To this end, for any $x \neq 0$, define $z = \|x\|^{-1}x$. Because A is linear, we have

$$\|x\|^{-1}\|Ax\| \;=\; \|Az\| \;\leqslant\; \|A\|_{\mathrm{op}}\|z\| \;=\; \|A\|_{\mathrm{op}}, \tag{0.1}$$

assuming the claim is true for z since $\|z\| = 1$. Multiplying by $\|x\|$ proves the validity of the reduction.

Now consider $x \in \mathbb{R}^n$ with $\|x\| = 1$. By definition, we have

$$\|Ax\| \;\leqslant\; \sup_{z\in\mathbb{R}^n:\|z\|=1} \|Az\| \;=\; \|A\|_{\mathrm{op}}, \tag{0.2}$$

which is what we want given $\|x\| = 1$.

(ii). Consider first $B = 0$ as a matrix. Then $\|AB\|_{\mathrm{op}} = 0$ since $AB = 0$ as a matrix, and the desired inequality is immediate.

For $B \neq 0$, we compute

$$\|AB\|_{\mathrm{op}} \;=\; \sup_{z\in\mathbb{R}^n:\|z\|=1} \|ABz\|. \tag{0.3}$$

Suppose $z \in \mathbb{R}^n$ such that $Bz = 0$. Then the quantity within the supremum is equal to 0. Thus, by part (i),

$$\begin{aligned}
\|AB\|_{\mathrm{op}} &= \sup_{z\in\mathbb{R}^n:\|z\|=1,Bz\neq 0} \|ABz\| && (0.4)\\
&= \sup_{z\in\mathbb{R}^n:\|z\|,Bz\neq 0} \frac{\|ABz\|}{\|Bz\|}\|Bz\| && (0.5)\\
&\leqslant \sup_{z\in\mathbb{R}^n:\|z\|,Bz\neq 0} \|A\frac{Bz}{\|Bz\|}\| \cdot \|B\|_{\mathrm{op}} && (0.6)\\
&\leqslant \|A\|_{\mathrm{op}}\|B\|_{\mathrm{op}}. && (0.7)
\end{aligned}$$

(iii). Given any $z = \sum_{i=1}^n z_i e_i$ with $\|z\| = \left(\sum_{i=1}^n |z_i|^2\right)^{\frac{1}{2}} = 1$, we have

$$\begin{aligned}
\|Az\| &= \left(\sum_{i=1}^n\left(\sum_{j=1}^n a_{ij}z_j\right)^2\right)^{\frac{1}{2}} && (0.8)\\
&\leqslant \left(\sum_{i=1}^n\left(\sum_{j=1}^n |a_{ij}|^2\right)\left(\sum_{j=1}^n |z_j|^2\right)\right)^{\frac{1}{2}} && (0.9)\\
&= \left(\sum_{i,j=1}^n |a_{ij}|^2\right)^{\frac{1}{2}}, && (0.10)
\end{aligned}$$

where the second line follows from Cauchy-Schwarz. Because this is uniform over $z \in \mathbb{R}^n$ with $\|z\| = 1$, we deduce the desired bound.

This inequality is not true in general. For example, if $A = \text{Id}$, so that $a_{ij} = \delta_{ij}$, then $\|A\|_{\text{op}} = 1$, and the other quantity is equal to $\sqrt{n} > 1$ if $n > 1$. If $n = 1$, then it's true, however.

(iv). Given any j, we first note that

$$[Ae_j]_i = a_{ij}. \tag{0.11}$$

We then have

$$\left(\sum_{i=1}^{n} |a_{ij}|^2\right)^{\frac{1}{2}} = \left\|Ae_j\right\| \tag{0.12}$$

$$\leqslant \|A\|_{\text{op}} \tag{0.13}$$

by part (i), since $\|e_j\| = 1$.

PROBLEM 2

We claim that we may assume $\|\cdot\|_2$ is the Euclidean norm. Indeed, if the claim were true for the Euclidean norm and any other norm $\|\cdot\|_1$, then for any two norms $\|\cdot\|_1, \|\cdot\|_3$, we have

$$c_1c_1'\|x\|_3 \leqslant c_1\|x\|_2 \leqslant \|x\|_1 \leqslant c_2\|x\|_2 \leqslant c_2c_2'\|x\|_3, \tag{0.14}$$

and also $c_1''\|x\|_3 \leqslant \|x\|_1 \leqslant c_2''\|x\|_3$ by the same argument.

For $x = 0$, the claim is certainly true. Moreover, we claim that it suffices to find constants c_1, c_2 such that the desired inequality is true for all x with $\|x\|_2 = 1$. Indeed, this reduction proceeds as Problem 1 part (i).

Thus, we are left with finding constants $c_1, c_2 > 0$ such that $c_1 \leqslant \|x\|_1 \leqslant c_2$ for all x with $\|x\|_2 = 1$. To this end, we first note $\{x : \|x\|_2 = 1\}$ is a compact subset of $\mathbb{R}^n$. Thus, it suffices to show that $x \mapsto \|x\|_1$ is continuous with respect to the Euclidean metric on $\mathbb{R}^n$, since continuous functions attain their minimum and maximum, and neither the minimum or maximum can be 0 or infinite if $\|x\|_2 = 1$.

To show the continuity, consider any $\varepsilon > 0$. For any $x = \sum_{i=1}^{n} x_i e_i$ and $y = \sum_{i=1}^{n} y_i e_i$, by the triangle inequality we have

$$|\|x\|_1 - \|y\|_1| \leqslant \|x - y\|_1 \tag{0.15}$$

$$= \left\|\sum_{i=1}^{n}(x_i - y_i)e_i\right\|_1 \tag{0.16}$$

$$\leqslant \sum_{i=1}^{n} |x_i - y_i|\|e_i\|_1 \tag{0.17}$$

$$\leqslant \|x - y\|_2 \sum_{i=1}^{n} \|e_i\|_1. \tag{0.18}$$

The summation is a constant depending only on n and the norm $\|\cdot\|_1$. Thus, if $\|x - y\|_2 < \frac{\varepsilon}{\sum_{i=1}^{n}\|e_i\|_1}$, we deduce $\|x - y\|_1 < \varepsilon$; this shows continuity, so we're done.

PROBLEM 3

(i). We let

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \tag{0.19}$$

We note that $A+B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, and that

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^k = \begin{cases} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & k \text{ odd} \\ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & k \text{ even}. \end{cases} \tag{0.20}$$

Moreover, we know $A^k = 0$ for all $k > 1$, and $B^k = 0$ for all $k > 1$. In particular, we compute

$$e^A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \tag{0.21}$$

$$e^B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \tag{0.22}$$

so that

$$e^A e^B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \tag{0.23}$$

$$= \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}. \tag{0.24}$$

Meanwhile,

$$e^{A+B} = \begin{pmatrix} \sum_{k=0}^{\infty} \frac{1}{(2k)!} & 0 \\ 0 & \sum_{k=0}^{\infty} \frac{1}{(2k!)} \end{pmatrix} + \begin{pmatrix} 0 & \sum_{k=0}^{\infty} \frac{1}{(2k+1)!} \\ \sum_{k=0}^{\infty} \frac{1}{(2k+1)!} & 0 \end{pmatrix}, \tag{0.25}$$

which is certainly not equal to $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$, since the diagonal entries are actually equal, for example.

(ii). The characteristic polynomial of $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ is equal to $p(\lambda) = \lambda^2$. Thus, for it to be diagonalizable, it must have two linearly independent vectors in the kernel. However, a vector is in the kernel if and only if $y = 0$ as seen by matrix multiplication. Thus, the kernel is given by the span of the first standard basis vector e_1, which shows the kernel is one-dimensional and thus the matrix is not diagonalizable.

Problem 4

(i). We compute

$$\text{Tr}(CD) = \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} d_{ji} \tag{0.26}$$

$$= \sum_{i,j=1}^{n} c_{ij} d_{ji} \tag{0.27}$$

$$= \sum_{j=1}^{n} \sum_{i=1}^{n} d_{ji} c_{ji} \tag{0.28}$$

$$= \text{Tr}(DC). \tag{0.29}$$

(ii). By part (i), letting $C = SA$ and $D = S^{-1}$, we know

$$\begin{aligned} \mathrm{Tr}(B) &= \mathrm{Tr}(SAS^{-1}) && (0.30)\\ &= \mathrm{Tr}(S^{-1}SA) && (0.31)\\ &= \mathrm{Tr}(A). && (0.32) \end{aligned}$$

(iii). Let $A = SDS^{-1}$, where D is an upper triangular matrix whose diagonal entries are the eigenvalues of A. Then part (ii) gives

$$\mathrm{Tr}(A) = \mathrm{Tr}(D) = \sum_{i=1}^{n} \lambda_i. \tag{0.33}$$

(iv). Suppose A is an upper triangular matrix. Then $\mathrm{e}^{\mathrm{Tr}(A)} = \prod_{i=1}^{n} \mathrm{e}^{a_{ii}}$. Moreover, for any $k \geqslant 0$, we know A^k is an upper triangular matrix whose diagonal entries are a_{ii}^k, respectively. Thus, e^A is also upper triangular has diagonal entries $\mathrm{e}^{a_{ii}}$, respectively just by definition of e^A and adding the resulting matrices and using the Taylor series for e^x. In particular, $\det \mathrm{e}^A$ is the product of the diagonal entries of e^A, which is $\prod_{i=1}^{n} \mathrm{e}^{a_{ii}}$. This completes the proof for A an upper triangular matrix.

In general, let $A = SBS^{-1}$ where B is upper triangular. Then

$$\begin{aligned} \mathrm{e}^A &= \mathrm{e}^{SBS^{-1}} && (0.34)\\ &= \sum_{k=0}^{\infty} \frac{(SBS^{-1})^k}{k!} && (0.35)\\ &= \sum_{k=0}^{\infty} SB^kS^{-1}\frac{1}{k!} && (0.36)\\ &= S\sum_{k=0}^{\infty} \frac{B^k}{k!}S^{-1} && (0.37)\\ &= S\mathrm{e}^B S^{-1}. && (0.38) \end{aligned}$$

Thus we know, given the upper triangular case,

$$\begin{aligned} \det \mathrm{e}^A &= \det(S\mathrm{e}^B S^{-1}) && (0.39)\\ &= \det \mathrm{e}^B && (0.40)\\ &= \prod_{i=1}^{n} \mathrm{e}^{b_{ii}}. && (0.41) \end{aligned}$$

On the other hand, we know $\mathrm{Tr}(A) = \mathrm{Tr}(B)$ by part (i), so that $\mathrm{e}^{\mathrm{Tr}(A)} = \mathrm{e}^{\mathrm{Tr}(B)} = \prod_{i=1}^{n} \mathrm{e}^{b_{ii}}$ as in the upper triangular case. This completes the proof.

PROBLEM 4

Let $A = \begin{pmatrix} 2 & 1 \\ 4 & -1 \end{pmatrix}$. The solution is given by

$$x(t) = \mathrm{e}^{tA}x(0) + \mathrm{e}^{tA}\int_0^t \mathrm{e}^{-sA}\begin{pmatrix} 0 \\ 5s \end{pmatrix} \mathrm{d}s. \tag{0.42}$$

Thus, we must compute e^{-sA} for all s. To this end, we diagonalize A:

- The characteristic polynomial of A is $p(\lambda) = (\lambda - 3)(\lambda + 2)$, so the eigenvalues of A are 3 and -2.
- An eigenvector for $\lambda = 3$ is $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. An eigenvector for $\lambda = -2$ is $\begin{pmatrix} 1 \\ -4 \end{pmatrix}$.

Thus, we have

$$A = Q\begin{pmatrix} 3 & 0 \\ 0 & -2 \end{pmatrix}Q^{-1}, \tag{0.43}$$

where $Q = \begin{pmatrix} 1 & 1 \\ 1 & -4 \end{pmatrix}$. In particular, we know, for all t,

$$\mathrm{e}^{tA} = Q\begin{pmatrix} \mathrm{e}^{3t} & 0 \\ 0 & \mathrm{e}^{-2t} \end{pmatrix}Q^{-1}. \tag{0.44}$$

This implies that e^{tA} has eigenvalue e^{3t} with eigenvector $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, and eigenvalue e^{-2t} with eigenvector $\begin{pmatrix} 1 \\ -4 \end{pmatrix}$.

To evaluate the integral, we first observe the projection

$$\begin{pmatrix} 0 \\ 5s \end{pmatrix} = s\begin{pmatrix} 1 \\ 1 \end{pmatrix} - s\begin{pmatrix} 1 \\ -4 \end{pmatrix}. \tag{0.45}$$

Thus, we know

$$\begin{aligned}
\mathrm{e}^{tA}\int_0^t \mathrm{e}^{-sA}\begin{pmatrix} 0 \\ 5s \end{pmatrix}\,\mathrm{d}s &= \int_0^t \mathrm{e}^{(t-s)A}\begin{pmatrix} 0 \\ 5s \end{pmatrix}\,\mathrm{d}s && (0.46)\\
&= \int_0^t \mathrm{e}^{3(t-s)}s\,\mathrm{d}s\cdot\begin{pmatrix} 1 \\ 1 \end{pmatrix} - \int_0^t \mathrm{e}^{-2(t-s)}s\,\mathrm{d}s\cdot\begin{pmatrix} 1 \\ -4 \end{pmatrix} && (0.47)\\
&= \frac{\mathrm{e}^{3t}-3t-1}{9}\cdot\begin{pmatrix} 1 \\ 1 \end{pmatrix} - \frac{\mathrm{e}^{-2t}+2t-1}{4}\cdot\begin{pmatrix} 1 \\ -4 \end{pmatrix}. && (0.48)
\end{aligned}$$

On the other hand, we have

$$\begin{aligned}
\mathrm{e}^{tA} &= \begin{pmatrix} 1 & 1 \\ 1 & -4 \end{pmatrix}\begin{pmatrix} \mathrm{e}^{3t} & 0 \\ 0 & \mathrm{e}^{-2t} \end{pmatrix}\begin{pmatrix} \frac{4}{5} & \frac{1}{5} \\ \frac{1}{5} & -\frac{1}{5} \end{pmatrix} && (0.49)\\
&= \begin{pmatrix} \mathrm{e}^{3t} & \mathrm{e}^{-2t} \\ \mathrm{e}^{3t} & -4\mathrm{e}^{-2t} \end{pmatrix}\begin{pmatrix} \frac{4}{5} & \frac{1}{5} \\ \frac{1}{5} & -\frac{1}{5} \end{pmatrix} && (0.50)\\
&= \begin{pmatrix} \frac{4}{5}\mathrm{e}^{3t}+\frac{1}{5}\mathrm{e}^{-2t} & \frac{1}{5}\mathrm{e}^{3t}-\frac{1}{5}\mathrm{e}^{-2t} \\ \frac{4}{5}\mathrm{e}^{3t}-\frac{4}{5}\mathrm{e}^{-2t} & \mathrm{e}^{3t}+\frac{4}{5}\mathrm{e}^{-2t} \end{pmatrix}. && (0.51)
\end{aligned}$$

Thus, our final answer is

$$x(t) = \begin{pmatrix} \frac{4}{5}\mathrm{e}^{3t}+\frac{1}{5}\mathrm{e}^{-2t} & \frac{1}{5}\mathrm{e}^{3t}-\frac{1}{5}\mathrm{e}^{-2t} \\ \frac{4}{5}\mathrm{e}^{3t}-\frac{4}{5}\mathrm{e}^{-2t} & \mathrm{e}^{3t}+\frac{4}{5}\mathrm{e}^{-2t} \end{pmatrix}x(0) + \frac{\mathrm{e}^{3t}-3t-1}{9}\cdot\begin{pmatrix} 1 \\ 1 \end{pmatrix} - \frac{\mathrm{e}^{-2t}+2t-1}{4}\cdot\begin{pmatrix} 1 \\ -4 \end{pmatrix}. \tag{0.52}$$

PROBLEM 5

Suppose that A is upper triangular. Then tA is upper triangular, so the eigenvalues of tA are the diagonal entries of tA. However, as in Problem 4 above, we know the e^{tA} is upper triangular with diagonal entries $\mathrm{e}^{ta_{ii}}$ if $a_{11}, \dots, a_{nn}$ are the diagonal entries of A. However, this implies $\mathrm{e}^{ta_{ii}}$ are the eigenvalues of e^{tA}, which completes the proof.

In general, as in Problem 4 above, we can write $A = SBS^{-1}$ where B is upper triangular, and the eigenvalues of B are exactly the eigenvalues of A. Similarly, we know $\mathrm{e}^{tA} = S\mathrm{e}^{tB}S^{-1}$, which means the eigenvalues of e^{tA} are the eigenvalues of e^{tB}. But now we just apply the case of upper triangular matrices.

PROBLEM 6

Given such an A, we write $A = D + N$, where D is diagonalizable and N is nilpotent, and D, N commute. Moreover, we know that on each generalized eigenspace $V^{(\lambda)}$ of A, both matrices D and N map $V^{(\lambda)}$ to iself; the matrix D acts by multiplication by λ, and N is nilpotent as a map on $V^{(\lambda)}$. Let $L > 0$ be such that $N^L = 0$ for all generalized eigenspaces $V^{(\lambda)}$.

For any $k > 0$, we have

$$A^k = (D+N)^k \tag{0.53}$$

$$= \sum_{\ell=0}^{k} \binom{k}{\ell} D^{k-\ell} N^{\ell}. \tag{0.54}$$

Indeed, this follows because D, N commute. For $k > L$, we have

$$A^k = \sum_{\ell=0}^{L-1} \binom{k}{\ell} N^{\ell} D^{k-\ell}. \tag{0.55}$$

Note that $\|N\|_{\mathrm{op}} < \infty$. To show that $A^k \to 0$, this is equivalent to showing $A^k v \to 0$ for all $v \in \mathbb{R}^n$. Moreover, it suffices to take v to be a vector in any generalized eigenspace $V^{(\lambda)}$ of A, since $\mathbb{R}^n$ is a direct sum of these generalized eigenspaces.

If $v \in V^{(\lambda)}$, we know $Dv = \lambda v$. Thus, we know

$$\|A^k v\| \leqslant \sum_{\ell=0}^{L-1} \binom{k}{\ell} \|N^{\ell} D^{k-\ell} v\| \tag{0.56}$$

$$\leqslant \sum_{\ell=0}^{L-1} \binom{k}{\ell} \|N\|_{\mathrm{op}}^{\ell} \|D^{k-\ell} v\| \tag{0.57}$$

$$\leqslant \sum_{\ell=0}^{L-1} \binom{k}{\ell} \|N\|_{\mathrm{op}}^{\ell} |\lambda|^{k-\ell} \|v\|. \tag{0.58}$$

Note $\binom{k}{\ell} \leqslant k^L$ for all $\ell = 0, \ldots, L-1$ and all k sufficiently large. Thus, we have

$$\sum_{\ell=0}^{L-1} \binom{k}{\ell} \|N\|_{\mathrm{op}}^{\ell} |\lambda|^{k-\ell} \|v\| \leqslant Lk^L \|N\|_{\mathrm{op}}^{\ell} |\lambda|^{k-\ell} \|v\|. \tag{0.59}$$

Because $|\lambda| < 1$ by assumption, we know the last quantity vanishes as $k \to \infty$, which completes the proof. Note that this also shows $\|D\|_{\mathrm{op}} < 1$ by the way.

PROBLEM 7

(i). We directly compute

$$\nabla \cdot F(x,y) = \partial_x F_1(x,y) + \partial_y F_2(x,y) \tag{0.60}$$

$$= \partial_x \partial_y \psi(x,y) - \partial_y \partial_x \psi(x,y) \tag{0.61}$$

$$= 0 \tag{0.62}$$

since ψ is smooth so its mixed partials agree.

(ii). It suffices to show that $\frac{\mathrm{d}}{\mathrm{d}t}\psi(X(t), Y(t)) = 0$ for all $t > 0$. To this end, we compute using the definition of F_1, F_2 and the ODEs for $X(t), Y(t)$:

$$\frac{\mathrm{d}}{\mathrm{d}t}\psi(X(t),Y(t)) = \partial_x \psi(X(t),Y(t))X'(t) + \partial_y \psi(X(t),Y(t))Y'(t) \tag{0.63}$$

$$= \partial_x \psi(X(t),Y(t))F_1(X(t),Y(t)) + \partial_y \psi(X(t),Y(t))F_2(X(t),Y(t)) \tag{0.64}$$

$$= \partial_x \psi(X(t),Y(t))\partial_y \psi(X(t),Y(t)) - \partial_y \psi(X(t),Y(t))\partial_x \psi(X(t),Y(t)) \tag{0.65}$$

$$= 0. \tag{0.66}$$

which completes the proof.

(iii). This is the same calculation as part (ii), except we differentiate $\frac{\mathrm{d}}{\mathrm{d}t}H(X(t),Y(t))$.