Norwegian University of Science and Technology
Department of Mathematical Sciences

TMA4145 Linear Methods
Fall 2018

Exercise set 8: Solutions

Note: When showing that $\|T\| = C$ for an operator T and a positive number C, we must show that $C = \sup\{\frac{\|Tx\|}{\|x\|} : x \neq 0\}$. One often divides the proof into two steps:

1. Show that $\|T\| \leq C$, often using a simple calculation. This means that C is some upper bound of $\{\frac{\|Tx\|}{\|x\|} : x \neq 0\}$.
2. Show that $\|T\| = C$ by showing that C is the *least* upper bound of $\{\frac{\|Tx\|}{\|x\|} : x \neq 0\}$.

These two steps should be used in problems 4 and 5.

1 Which of the following transformations are linear?

a) $T : P_2(\mathbb{R}) \to P_3(\mathbb{R})$ defined by $T(p)(x) = xp(x) + p'(x)$, where $P_n(\mathbb{R})$ denotes the vector space of real-valued polynomials of degree at most n.

b) $T : \mathbb{C}^2 \to \mathbb{C}^2$ defined by $T(z_1, z_2) = (\overline{z_1}, \overline{z_2})$, where $\mathbb{C}^2$ is a vector space over $\mathbb{R}$.

Does the conclusion change if $\mathbb{C}^2$ is considered as a vector space over $\mathbb{C}$? Explain.

c) Let $M_{n\times n}(\mathbb{R})$ denote the space of all $n \times n$ matrices with real entries.

i) $T : M_{n\times n}(\mathbb{R}) \to M_{n\times n}(\mathbb{R})$, $T(A) = A^2$.

ii) $T : M_{n\times n}(\mathbb{R}) \to \mathbb{R}$, $T(A) = \det A$.

Solution. a) T is linear over $\mathbb{R}$. If $\lambda, \mu \in \mathbb{R}$ and $p, q \in P_2$, then

$$\begin{aligned} T(\lambda p + \mu q)(x) &= x\left[\lambda p(x) + \mu q(x)\right] + \lambda p'(x) + \mu q'(x) \\ &= \lambda x p(x) + \lambda p'(x) + \mu x q(x) + \mu q'(x) \\ &= \lambda[xp(x) + p'(x)] + \mu[xq(x) + q'(x)] \\ &= \lambda T(p)(x) + \mu T(q)(x). \end{aligned}$$

We have used that differentiation is linear: $(\lambda p + \mu q)' = \lambda p' + \mu q'$.

b) T is linear when $\mathbb{C}$ is considered a vector space over $\mathbb{R}$. Let $\lambda, \mu \in \mathbb{R}$ and $(z_1, z_2), (w_1, w_2) \in \mathbb{C}^2$. By definition of scalar multiplication and addition in $\mathbb{C}^2$, we have that $\lambda(w_1, w_2) + \mu(z_1, z_2) = (\lambda w_1 + \mu z_1, \lambda w_2 + \mu z_2)$. Therefore

$$\begin{aligned}
T(\lambda(w_1, w_2) + \mu(z_1, z_2)) &= T(\lambda w_1 + \mu z_1, \lambda w_2 + \mu z_2) \\
&= (\overline{\lambda w_1 + \mu z_1}, \overline{\lambda w_2 + \mu z_2}) \\
&= (\lambda \overline{w_1} + \mu \overline{z_1}, \lambda \overline{w_2} + \mu \overline{z_2}) \quad (\overline{\lambda} = \lambda, \overline{\mu} = \mu \text{ since } \lambda, \mu \in \mathbb{R}) \\
&= (\lambda \overline{w_1}, \lambda \overline{w_2}) + (\mu \overline{z_1}, \mu \overline{z_2}) \\
&= \lambda(\overline{w_1}, \overline{w_2}) + \mu(\overline{z_1}, \overline{z_2}) \\
&= \lambda T(w_1, w_2) + \mu T(z_1, z_2).
\end{aligned}$$

If we consider $\mathbb{C}$ to be a vector space over $\mathbb{C}$, then T is not linear. For instance, we have that

$$\begin{aligned}
T(i(1,0)) &= T(i, 0) \\
&= (\bar{i}, 0) \\
&= (-i, 0) \\
&= -i(1, 0),
\end{aligned}$$

which does not equal $iT(1,0) = i(1,0)$.

c) i) T is not linear. In general, multiplication of matrices satisfies, for $A, B \in M_{n\times n}$, that

$$\begin{aligned}
T(A + B) &= (A + B)(A + B) \\
&= A^2 + AB + BA + B^2,
\end{aligned}$$

which in general does not equal $T(A) + T(B) = A^2 + B^2$. To obtain a concrete counterexample, let $A = B = I_n$ – the identity matrix in $M_{n\times n}$. Then $T(I_n + I_n) = I_n^2 + I_n I_n + I_n I_n + I_n^2 = 4I_n$ since $I_n^2 = I_n$. On the other hand, $T(I_n) + T(I_n) = I_n^2 + I_n^2 = 2I_n$. Therefore $T(I_n + I_n) \neq T(I_n) + T(I_n)$.
ii) T is only linear when $n = 1$. From linear algebra, you know that the determinant function satisfies

$$\det(\lambda A) = \lambda^n \det(A)$$

. So if $n \neq 1$, we do *not* have $T(\lambda A) = \lambda T(A)$, hence T is not linear. For $n = 1$ the determinant is simply $det(a) = a$ for $a \in M_{1\times 1}(\mathbb{R}) = \mathbb{R}$, i.e. the identity transformation, and the identity transformation is linear.

$\boxed{2}$ Let X and Y be normed spaces. Show that a linear map $T : X \to Y$ is not continuous if and only if there exists a sequence of unit vectors (x_n) in X such that $\|Tx_n\| \geq n$ for $n \in \mathbb{N}$.

Solution. Since T is a linear mapping between normed spaces, it is bounded if and only if it is continuous. Assume that T is not bounded. Then by lemma 4.4 we have that

$$\sup\{\|Tx\|_Y : \|x\|_X = 1\} = \infty.$$

Hence no $n \in \mathbb{N}$ is an upper bound for the set $\{\|Tx\|_Y : \|x\|_X = 1\}$, which means that for every $n \in \mathbb{N}$ there exists some $x_n \in X$ with $\|x_n\|_X = 1$ and $\|Tx_n\|_Y > n$. This defines a sequence (x_n) with the desired properties.
Conversely, assume that we have a sequence (x_n) of unit vectors with $\|Tx_n\| \geq n$ for every $n \in \mathbb{N}$. Since the supremum is an upper bound, we must then have

$$\sup\{\|Tx\|_Y : \|x\|_X = 1\} \geq n \text{ for every } n \in \mathbb{N}.$$

Hence $\sup\{\|Tx\|_Y : \|x\|_X = 1\} = \infty$, so the operator is not bounded.

$\boxed{3}$ Let X and Y be vector spaces, both real or both complex. Let $T : X \to Y$ be a linear operator with some range $\operatorname{ran}(T) \subset Y$. Show that:

a) The inverse operator $T^{-1} : \operatorname{ran}(T) \to X$ exists if and only if

$$Tx = 0 \quad \Rightarrow \quad x = 0.$$

(In other words: if and only if $\ker(T) = \{0\}$.)

b) If T^{-1} exists, it is a linear operator.

c) Even if T is a bounded operator, its inverse T^{-1} need not be.

Note: An inverse operator $T^{-1} : \operatorname{ran}(T) \to X$ is an operator satisfying $T^{-1}(T(x)) = x$ and $T(T^{-1}(y))$ for any $x \in X$ and $y \in \operatorname{ran}(T)$.

Solution. a) First assume that the inverse $T^{-1} : \operatorname{ran}(T) \to X$ exists. Note that since $T(0) = 0$ by linearity of T, we have that $0 \in \operatorname{ran}(T)$ and $T^{-1}(0) = 0$. If $T(x) = 0$, then we can apply T^{-1} to both sides of the equation to get that

$$T^{-1}(T(x)) = T^{-1}(0)$$
$$x = 0$$

since $T^{-1}(T(x)) = x$ by the definition of being an inverse operator.

Then assume that $T(x) = 0 \implies x = 0$. Since every element of $\operatorname{ran}(T)$ is of the form $T(x)$ for some $x \in X$ (this is the definition of $\operatorname{ran}(T)$), we would like to define a candidate for the inverse operator T^{-1} by

$$T^{-1}(T(x)) = x.$$

However, this operator might not be well-defined – what if we had two points $x, z \in X$ such that $T(x) = T(z)$, should then $T^{-1}(T(x)) = T^{-1}(T(z))$ be x or z? Luckily, our extra assumption makes sure that this cannot happen. If $T(x) = T(z)$, then

$T(x-z)=0$ by linearity, hence $x-z=0$ by our assumption. Therefore $x=z$. It follows that T^{-1} is well-defined. One then easily checks that T^{-1} is an inverse operator. By definition $T^{-1}(T(x))=x$ for $x\in X$, and for $y=T(x)\in \operatorname{ran}(T)$ we have

$$T(T^{-1}(y))=T(T^{-1}(Tx))=T(x)=y.$$

b) Let λ,μ be scalars (real or complex, depending on X and Y), and let $y_1,y_2\in \operatorname{ran}(T)$. By definition of $\operatorname{ran}(T)$, we can find $x_1,x_2\in X$ such that $T(x_1)=y_1$ and $T(x_2)=y_2$. Then

$$\begin{aligned}T^{-1}(\lambda y_1+\mu y_2)&=T^{-1}(\lambda T(x_1)+\mu T(x_2))\\&=T^{-1}(T(\lambda x_1+\mu x_2))\quad \text{since } T \text{ is linear}\\&=\lambda x_1+\mu x_2\quad \text{by definition of inverse operator}\\&=\lambda T^{-1}(y_1)+\mu T^{-1}(y_2)\quad \text{since } T(x_1)=y_1, T(x_2)=y_2.\end{aligned}$$

Hence T^{-1} is linear.

c) Let X be the space of real-valued sequences with only finitely many non-zero elements; X consists of sequences of the form

$$x=(x_1,x_2,\dots,x_n,0,0,\dots).$$

X is a normed space with the norm from ℓ^∞, namely

$$\|x\|=\sup_{n\in\mathbb{N}}|x_n|.$$

Let $T:X\to X$ be the linear map given by

$$T(x)=(x_1,\frac{1}{2}x_2,\frac{1}{3}x_3,\dots,\frac{1}{n}x_n,0,0,\dots).$$

T is bounded on X since

$$\begin{aligned}\|T(x)\|&=\sup_{n\in\mathbb{N}}|\frac{1}{n}x_n|\\&\le\sup_{n\in\mathbb{N}}|x_n|\\&=\|x\|.\end{aligned}$$

It is not much work to show that $\operatorname{ran}(T)=X$, and that the inverse is given by

$$T^{-1}(y)=(y_1,2y_2,3y_3,\dots,ny_n,0,0,\dots.)$$

We will use problem 2 to show that T^{-1} is not bounded. Let (x^n) be the sequence in X given by

$$x^n=(0,0,\dots,0,1,0,\dots),$$

where the only non-zero element appears in position n. Then $\|x^n\|=1$ for each $n\in\mathbb{N}$, yet

$$T^{-1}(x^n)=(0,0,\dots,0,n,0,\dots),$$

so $\|T^{-1}(x^n)\|=n$. So we may conclude using problem 2 that T^{-1} is not bounded.

$\boxed{4}$ Let T be a linear mapping $T : (\mathbb{R}^n, \|.\|_\infty) \to (\mathbb{R}^n, \|.\|_\infty)$ given by a $n \times n$ matrix A. Show that the operator norm of T in terms of A is given by $\|T\| = \max_{i=1,\dots,n} \sum_{j=1}^n |a_{ij}|$.

Solution.
Please note that there is an example at the end of the solution, to illustrate what we are doing. Let $x = (x_1, x_2, ..., x_n) \in \mathbb{R}^n$. Then the vector $Ax = ((Ax)_1, (Ax)_2, ...(Ax)_n)$ is given by

$$(Ax)_i = \sum_{j=1}^n a_{ij}x_j,$$

by the definition of the product of a matrix and a vector. Hence

$$\|Ax\|_\infty = \max_{i=1,\dots,n} |\sum_{j=1}^n a_{ij}x_j| \leq \max_{i=1,\dots,n} \sum_{j=1}^n |a_{ij}||x_j| \leq \|x\|_\infty \max_{i=1,\dots,n} \sum_{j=1}^n |a_{ij}|,$$

where we have used the triangle inequality and the fact that $|x_i| \leq \|x\|_\infty$ for any $1 \leq i \leq n$. This shows that

$$\frac{\|Tx\|_\infty}{\|x\|_\infty} \leq \max_{i=1,\dots,n} \sum_{j=1}^n |a_{ij}| \quad \text{for } x \neq 0, \tag{1}$$

and hence $\|T\| \leq \max_{i=1,\dots,n} \sum_{j=1}^n |a_{ij}|$ since $\|T\| = \sup\{\frac{\|Tx\|_\infty}{\|x\|_\infty} : x \neq 0\}$.
One way to show that $\|T\| = \max_{i=1,\dots,n} \sum_{j=1}^n |a_{ij}|$ would be to find some vector $x \in \mathbb{R}^n$ such that

$$\frac{\|Tx\|_\infty}{\|x\|_\infty} \geq \max_{i=1,\dots,n} \sum_{j=1}^n |a_{ij}|. \tag{2}$$

This would show that $\max_{i=1,\dots,n} \sum_{j=1}^n |a_{ij}|$ is not only an upper bound for $\{\frac{\|Tx\|_\infty}{\|x\|_\infty} : x \neq 0\}$, as we saw in equation (1), but actually the *least* upper bound. Let us therefore look for an x satisfying (2).
Assume that the maximal row sum is achieved for the index k – meaning that $\max_{i=1,\dots,n} \sum_{j=1}^n |a_{ij}| = \sum_{j=1}^n |a_{kj}|$. We are looking for $x \in \mathbb{R}^n$ such that

$$\frac{\|Tx\|_\infty}{\|x\|_\infty} \geq \sum_{j=1}^n |a_{kj}|.$$

Define $x = (x_1, x_2, ..., x_n) \in \mathbb{R}^n$ by

$$x_j = \begin{cases} 1 & \text{if } a_{kj} > 0 \\ 0 & \text{if } a_{kj} = 0 \\ -1 & \text{if } a_{kj} < 0. \end{cases}$$

In other words, we construct x from the row $(a_{k1}, a_{k2}, ..., a_{kn})$ of A by defining the j'th element of x to be the *sign* of the j'th element of this row. This means that $x_j a_{kj} = |a_{kj}|$ for every $j = 1, 2, ...n$. We then get that

$$(Tx)_k = (Ax)_k = \sum_{j=1}^n a_{kj}x_j = \sum_{j=1}^n |a_{kj}|,$$

where the last equality follows from the way we defined x. This means that $\|Tx\|_\infty = \sum_{j=1}^n |a_{kj}|$ by the definition of the ∞-norm. Clearly $\|x\|_\infty = 1$ [1]. Hence

$$\frac{\|Tx\|_\infty}{\|x\|_\infty} \geq \sum_{j=1}^n |a_{kj}|,$$

which is what we needed to show.

Example *By no means expected as part of a solution – for pedagogical reasons only.* Let A be defined by

$$A = \begin{pmatrix} 2 & -3 & 1 \\ 0 & 2 & 2 \\ 3 & 0 & -5. \end{pmatrix}$$

In this case $n = 3$, and we would like to understand what the expression $\max_{i=1,2,3} \sum_{j=1}^3 |a_{ij}|$ means in this case. Note that i denotes the row number, so for a fixed i the sum $\sum_{j=1}^3 |a_{ij}|$ is the sum of the absolute value of row i. In our case:

- For $i = 1$, the sum is $2 + |-3| + 1 = 6$.
- For $i = 2$, the sum is $0 + 2 + 2 = 4$.
- For $i = 3$, the sum is $3 + 0 + |-5| = 8$.

The expression $\max_{i=1,2,3} \sum_{j=1}^3 |a_{ij}|$ is the maximum of these three numbers, and clearly the maximum is achieved in the third row and equals 8. The second part of the solution constructed some vector x such that $\|Ax\|_\infty \geq \max_{i=1,2,3} \sum_{j=1}^3 |a_{ij}|$. In our case we know that the last expression actually equals 8, and is the sum of the absolute values of the elements in the third row. How can we then find some unit vector $x \in \mathbb{R}^3$ such that $\|Ax\|_\infty \geq 8$? We can do this by noting that element 3 in the vector Ax is the dot product between the third row of A and x:

$$Ax = \begin{pmatrix} 2 & -3 & 1 \\ 0 & 2 & 2 \\ 3 & 0 & -5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} ? \\ ? \\ 3x_1 + 0x_2 - 5x_3 \end{pmatrix}$$

So in order to make the third element of Ax be equal to 8, we could obviously choose $x = (1, 0, -1)$. Then $\|Ax\|_\infty \geq 8$, and it was such a vector x that we needed to construct. This is the reasoning used to construct the solution in the general case.

$\boxed{5}$ Let T be the integral operator $Tf(x) = \int_0^1 k(x,y)f(y)dy$ defined by a kernel $k \in C([0,1] \times [0,1])$ such that $k(x,y) \geq 0$ for any $(x,y) \in [0,1] \times [0,1]$. Show that the operator norm of T as a mapping on $C[0,1]$ with respect to $\|.\|_\infty$-norm is $\|T\| = \max_{x\in[0,1]} \int_0^1 |k(x,y)|dy$.

[1] We could in theory also have that $\|x\|_\infty = 0$, but this means that each $a_{kj} = 0$, and hence $\|T\| = 0$ since $\|T\| = \sum_{j=1}^n |a_{kj}|$.

Solution. Let $f \in C[0,1]$. We calculate using the triangle inequality for integrals that

$$\begin{aligned}\|Tf\|_\infty &= \max_{x\in[0,1]} \left|\int_0^1 k(x,y)f(y)dy\right| \\ &\le \max_{x\in[0,1]} \int_0^1 |k(x,y)||f(y)|dy \\ &\le \|f\|_\infty \max_{x\in[0,1]} \int_0^1 |k(x,y)|dy \\ &= \|f\|_\infty \max_{x\in[0,1]} \int_0^1 |k(x,y)|dy.\end{aligned}$$

As we did in problem 4, we may conclude from this inequality that $\|T\| \le \max_{x\in[0,1]} \int_0^1 |k(x,y)|dy$. We would now like to show that $\|T\| = \max_{x\in[0,1]} \int_0^1 |k(x,y)|dy$, and similarly to problem 4 it will be enough to find some function $f \in C[0,1]$ such that $\frac{\|Tf\|_\infty}{\|f\|_\infty} \ge \max_{x\in[0,1]} \int_0^1 |k(x,y)|dy$. Assume that the integral $\int_0^1 |k(x,y)|dy$ attains its maximum at the point x'– meaning that $\max_{x\in[0,1]} \int_0^1 |k(x,y)|dy = \int_0^1 |k(x',y)|dy$. If we pick f to be the constant function $f(x) = 1$, then

$$Tf(x') = \int_0^1 k(x',y)dy = \int_0^1 |k(x',y)|dy,$$

which implies that $\|Tf\|_\infty \ge \int_0^1 |k(x',y)|f(y)dy$, and hence

$$\frac{\|Tf\|_\infty}{\|f\|_\infty} \ge \int_0^1 |k(x',y)|dy = \max_{x\in[0,1]} \int_0^1 |k(x,y)|dy,$$

which is what we needed to show.
Note: We have used that k is positive to get that $|k(x,y)| = k(x,y)$ for $x, y \in \mathbb{R}$. If this was not true, we would not get that $\|Tf\|_\infty = \int_0^1 |k(x',y)|f(y)dy$ when f is the constant function 1. The obvious way of fixing this would be to follow in the steps of problem (3), and replace this constant function with f defined by

$$f(y) = \begin{cases} 1 & \text{if } k(x',y) > 0 \\ 0 & \text{if } k(x',y) = 0 \\ -1 & \text{if } k(x',y) < 0. \end{cases}$$

The problem with this procedure is that this function is not continuous in general, which is a big problem since we are working in the space of continuous functions! Of course, if k is a positive function, this f will simply be the constant function 1. There are two ways of getting rid of this problem:

1. By a slightly more technical proof, it is possible to show that $\|T\| = \max_{x\in[0,1]} \int_0^1 |k(x,y)|dy$ even when k is not positive.

2. One could also change the setting of the problem by considering T to be an operator from $L^\infty([0,1])$ to $L^\infty([0,1])$. Here $L^\infty([0,1])$ is the much larger space of all bounded functions on $[0,1]$ – they don't have to be continuous. You

will meet this space later in the course. In particular, the f that we defined by taking the sign of $k(x', y)$ actually belongs to $L^\infty([0,1])$, since its absolute value never exceeds 1. Hence we may apply T to this f, and we would find that

$$\|Tf\|_\infty = \int_0^1 |k(x', y)dy.$$