Differential Equations (Math 285)

H26 Let (M, d) be a metric space and $(a, b) \in M \times M$.

a) Show that the metric d is *continuous* in the following sense:
For every $\epsilon > 0$ there exists $\delta > 0$ such that $d(x, a) < \delta \wedge d(y, b) < \delta$ implies $|d(x, y) - d(a, b)| < \epsilon$.

Hint: First derive the so-called *quadrangle inequality* $|d(x, y) - d(a, b)| \leq d(x, a) + d(y, b)$.

b) Using a), show in detail that $x_n \to a$ and $y_n \to b$ implies $d(x_n, y_n) \to d(a, b)$. (A special case of this, viz. $d(x_n, b) \to d(a, b)$, is used in the proof of Part (2) of Banach's Fixed-Point Theorem.)

H27 a) Show that a closed subset N of a complete metric space (M, d) is complete in the induced metric $d|_N \colon N \times N \to \mathbb{R}$, $(x, y) \mapsto d(x, y)$.

b) Conversely, show that a subset of a metric space that is complete in the induced metric must be closed.

H28 a) Compute the norms $\|\mathbf{A}\|$ of the following matrices $\mathbf{A} \in \mathbb{R}^{2\times 2}$ directly from the definition and compare them with their Frobenius norms $\|\mathbf{A}\|_{\mathrm{F}}$.

$$\begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}, \quad \begin{pmatrix} 2 & 0 \\ 0 & -3 \end{pmatrix}, \quad \begin{pmatrix} \frac{1}{2} & \pm 1 \\ 0 & \frac{1}{2} \end{pmatrix}.$$

b) Show that the norm $\|\mathbf{D}\|$ of a diagonal matrix $\mathbf{D} \in \mathbb{R}^{n\times n}$ is the largest absolute value of an entry on the diagonal.

c) Show that if $T \colon \mathbb{R}^n \to \mathbb{R}^n$ is continuously differentiable with $\mathbf{J}_T(\mathbf{x}^*) = \mathbf{0}$, there exists $r > 0$ such that $T \colon \overline{\mathrm{B}_r(\mathbf{x}^*)} \to \overline{\mathrm{B}_r(\mathbf{x}^*)}$ forms a contraction with constant $C = \frac{1}{2}$.

d) Give an example of a linear map $T \colon \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{x} \mapsto \mathbf{A}\mathbf{x}$ that forms a contraction but has Frobenius norm > 1.

H29 Solve the initial value problem

$$y'' + |y| = 0, \quad y(0) = 0, \; y'(0) = 1.$$

Your solution should have the (maximal) domain $\mathbb{R}$. Is the solution unique?
Hint: The solution of Example 10 in `lecture1-3_handout.pdf` and Exercise H18 a) of Homework 3 may help.

H30 *Optional Exercise*

Let M be a set and $d\colon M \times M \to \mathbb{R}$ a function satisfying $d(a,a) = 0$ for $a \in M$, $d(a,b) \neq 0$ for $a, b \in M$ with $a \neq b$, and $d(a,b) \leq d(b,c)) + d(c,a)$ for $a, b, c \in M$.

a) Show that d is a metric.

b) Does this conclusion also hold if $d(a,b) \leq d(b,c)) + d(c,a)$ is replaced by the ordinary triangle inequality $d(a,b) \leq d(a,c)) + d(c,b)$?

H31 *Optional Exercise*

a) Prove that $\mathbb{R}^{n\times n} \to \mathbb{R}$, $\mathbf{A} \mapsto \|\mathbf{A}\|$ satisfies (N1)–(N4).

b) Repeat a) for the Frobenius norm $\mathbb{R}^{n\times n} \to \mathbb{R}$, $\mathbf{A} \mapsto \|\mathbf{A}\|_{\mathrm{F}}$.

c) Show that $\|\mathbf{A}\| \leq \|\mathbf{A}\|_{\mathrm{F}}$ for all matrices $\mathbf{A} \in \mathbb{R}^{n\times n}$ or, equivalently, $|\mathbf{A}\mathbf{x}| \leq \|\mathbf{A}\|_{\mathrm{F}} \, |\mathbf{x}|$ for all $\mathbf{A} \in \mathbb{R}^{n\times n}$ and $\mathbf{x} \in \mathbb{R}^n$.
Hint: Use $\|\mathbf{A}\| = \max\bigl\{|\mathbf{A}\mathbf{x}|\,;\mathbf{x} \in \mathbb{R}^n, |\mathbf{x}| = 1\bigr\}$ and the Cauchy-Schwarz Inequality for vectors in $\mathbb{R}^n$.

d) For $\mathbf{A} \in \mathbb{R}^{n\times n}$ show $\|\mathbf{A}\| = \bigl\|\mathbf{A}^{\mathsf{T}}\bigr\|$.

e) Suppose $\mathbf{A} \in \mathbb{R}^{n\times n}$ is invertible, $\mathbf{B} = \mathbf{A}^{\mathsf{T}}\mathbf{A}$, and $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ are the eigenvalues of $\mathbf{B}$. Show $\|\mathbf{A}^{-1}\| = 1/\sqrt{\lambda_n}$.

f) Using the notation in e), show $\|A\|_{\mathrm{F}} = \sqrt{\lambda_1 + \lambda_2 + \cdots + \lambda_n}$. (This yields an alternative proof of the inequality $\|\mathbf{A}\| = \sqrt{\lambda_1} \leq \|\mathbf{A}\|_{\mathrm{F}}$.)
Hint: $\|\mathbf{A}\|_{\mathrm{F}} = \mathrm{tr}\bigl(\mathbf{A}^{\mathsf{T}}\mathbf{A}\bigr)$; cf. Math257 in Fall 2022, Exercise H58 a) of Homework 11.

Due on Fri Mar 24, 4 pm

Matrix norms (required for H28) will be discussed in the lecture on Mon Mar 20. Exercises H28 c), d) are also optional. The optional exercises can be handed in until Fri Mar 31, 4 pm.

Solutions

26 a) Applying the triangle inequality twice, we have

$$\begin{aligned} d(x,y) &\leq d(x,a)+d(a,y) \\ &\leq d(x,a)+d(a,b)+d(b,y). \\ \implies d(x,y)-d(a,b) &\leq d(x,a)+d(y,b) \end{aligned}$$

Interchanging x, a as well as y, b in this inequality turns the left-hand side into $d(a,b) - d(x,y)$ and preserves the right-hand side, so that we also have $d(a,b) - d(x,y) \leq d(x,a) + d(y,b)$. Thus $\pm\big(d(x,y) - d(a,b)\big) \leq d(x,a) + d(y,b)$, which is equivalent to the quadrangle inequality.

With the quadrangle inequality at hand the continuity of d is easy to prove: Just choose $\delta = \epsilon/2$ as response to ϵ.

b) Let $\epsilon > 0$ be given. There exists $N_1 \in \mathbb{N}$ such that $d(x_n, a) < \epsilon/2$ for all $n > N_1$, and $N_2 \in \mathbb{N}$ such that $d(y_n, b) < \epsilon/2$ for all $n > N_2$. Using the qudrangle inequality, we then have

$$|d(x_n, y_n) - d(a,b)| \leq d(x_n, a) + d(y_n, b) < \epsilon/2 + \epsilon/2 = \epsilon$$

for all $n \geq \max\{N_1, N_2\}$. Thus $N = \max\{N_1, N_2\}$ can serve as response to ϵ in a proof of $d(x_n, y_n) \to d(a,b)$.

27 a) If (x_n) is a Cauchy sequence in N, it is a fortiori a Cauchy sequence in M and hence converges to some $a \in M$, since (M,d) is complete. But "N closed" means that N contains all limit points of sequences in N, so $a \in N$ and (x_n) converges in $(N, d|_N)$, which is therefore complete as well.

b) Using the notation in a), let (x_n) be a sequence in N, which has a limit in M, say a. Then (x_n) must be a Cauchy sequence, and hence convergent in N, since $(N, d|_N)$ is complete. Since limits of sequences are unique (the easily proved analogue for metric spaces of Exercise W18 a) of Worksheet 6 in Calculus III, Fall 2022), this implies $a \in N$. Thus N contains all limit points of sequences in N and hence is closed.

Remarks: By the term "limit point" I mean just "limit", but some people would interpret "limit points" as "accumulation points" of not necessarily convergent sequences. In fact, since closed subsets are also characterized as subsets containing all their accumulation points, both views are admitted for this exercise.

Note that in b) the completeness of M is not required, and hence b) holds also for complete subspaces of incomplete metric spaces.

28 a) Set $\mathbf{x} = \begin{pmatrix} \sin x & \cos x \end{pmatrix}^{\mathrm{T}}$ for $x \in [0, 2\pi)$.
In what follows, all maxima are taken over $x \in [0, 2\pi)$ (or over $\mathbb{R}$, which amounts to the same).

i) The norms of $\mathbf{A} = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}$ are shown below.

$$\|\mathbf{A}\| = \max\{|\mathbf{A}\mathbf{x}|\} = \max\left\{\left|\begin{pmatrix} 2\sin x + 2\cos x \\ 2\sin x + 2\cos x \end{pmatrix}\right|\right\} = \max\left\{\sqrt{2\,(2\sin x + 2\cos x)^2}\right\} = 4$$

$$\|\mathbf{A}\|_F = \sqrt{2^2 + 2^2 + 2^2 + 2^2} = 4$$

Therefore for $\mathbf{A} = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}$, $\|\mathbf{A}\| = \|\mathbf{A}\|_F$.

ii) The norms of $\mathbf{A} = \begin{pmatrix} 2 & 0 \\ 0 & -3 \end{pmatrix}$ are shown below.

$$\|\mathbf{A}\| = \max\{|\mathbf{Ax}|\} = \max\left\{\left|\begin{pmatrix} 2\sin x \\ -3\cos x \end{pmatrix}\right|\right\} = \max\left\{\sqrt{2(\sin x)^2 + (-3\cos x)^2}\right\} = 3$$
$$\|\mathbf{A}\|_F = \sqrt{2^2 + 0^2 + 0^2 + 3^2} = \sqrt{13}$$

Therefore for $\mathbf{A} = \begin{pmatrix} 2 & 0 \\ 0 & -3 \end{pmatrix}$, $\|\mathbf{A}\| < \|\mathbf{A}\|_F$.

iii) The norms of $\mathbf{A} = \begin{pmatrix} \frac{1}{2} & \pm 1 \\ 0 & \frac{1}{2} \end{pmatrix}$ are shown below.

$$\begin{aligned}\|\mathbf{A}\| &= \max\{|\mathbf{Ax}|\} = \max\left\{\left|\begin{pmatrix} \frac{1}{2}\sin x \pm \cos x \\ \frac{1}{2}\cos x \end{pmatrix}\right|\right\} \\ &= \max\left\{\sqrt{\left(\frac{1}{2}\sin x \pm \cos x\right)^2 + \left(\frac{1}{2}\cos x\right)^2}\right\} = \sqrt{\frac{3}{4} + \frac{\sqrt{2}}{2}} = \frac{1+\sqrt{2}}{2} \approx 1.207,\end{aligned}$$
$$\|\mathbf{A}\|_F = \sqrt{\frac{1}{2}^2 + 1^2 + 0^2 + \frac{1}{2}^2} = \frac{\sqrt{6}}{2} \approx 1.225$$

(For the former, using the Calculus I machinery one finds that $x \mapsto \left(\frac{1}{2}\sin x \pm \cos x\right)^2 + \left(\frac{1}{2}\cos x\right)^2 = \frac{1}{4} + \cos^2 x \pm \sin x \cos x$ is maximized at $x_1 = \pm\pi/8$ and $x_2 = \pm 5\pi/8$ with value $\frac{3}{4} + \frac{1}{2}\sqrt{2} = \frac{3+2\sqrt{2}}{4}$.)
Therefore for $\mathbf{A} = \begin{pmatrix} \frac{1}{2} & \pm 1 \\ 0 & \frac{1}{2} \end{pmatrix}$, $\|\mathbf{A}\| < \|\mathbf{A}\|_F$.
Remark: Here is the alternative computation using the formula $\|\mathbf{A}\| = \sqrt{\rho(\mathbf{A}^\mathsf{T}\mathbf{A})}$:

$$\begin{aligned}\mathbf{B} &= \mathbf{A}^\mathsf{T}\mathbf{A} = \begin{pmatrix} \frac{1}{4} & \pm\frac{1}{2} \\ \pm\frac{1}{2} & \frac{5}{4} \end{pmatrix}, \\ \chi_\mathbf{B}(X) &= X^2 - \frac{3}{2}X + \frac{5}{16} - \frac{1}{4} = X^2 - \frac{3}{2}X + \frac{1}{16}, \\ \lambda_{1/2} &= \frac{1}{2}\left(\frac{3}{2} \pm \sqrt{\frac{9}{4} - \frac{4}{16}}\right) = \frac{1}{4}\left(3 \pm 2\sqrt{2}\right) = \frac{1}{4}\left(1 \pm \sqrt{2}\right)^2, \\ \rho(\mathbf{B}) &= \lambda_1 = \frac{1}{4}\left(1+\sqrt{2}\right)^2, \\ \|\mathbf{A}\| &= \sqrt{\lambda_1} = \frac{1}{2}\left(1+\sqrt{2}\right).\end{aligned}$$

b) Suppose $\mathbf{D}$ has diagonal entries $d_1, \dots, d_n$, and assume w.l.o.g. that $|d_1| \geq |d_i|$ for $2 \leq i \leq n$.
For $\mathbf{e}_1 = (1, 0, \dots, 0)^\mathsf{T}$ we have $\mathbf{De}_1 = (d_1, 0, \dots, 0)^\mathsf{T}$ and hence $\frac{|\mathbf{De}_1|}{|\mathbf{e}_1|} = \frac{|d_1|}{1} = |d_1|$. This shows $\|\mathbf{D}\| \geq |d_1|$.

Now consider an arbitrary nonzero (column) vector $\mathbf{x} = (x_1, \dots, x_n)^\mathsf{T}$ in $\mathbb{R}^n$. Then

$$\begin{aligned}\mathbf{Dx} &= (d_1x_1, \dots, d_nx_n)^\mathsf{T},\\ |\mathbf{Dx}|^2 &= d_1^2x_1^2 + \cdots + d_n^2x_n^2 \le d_1^2(x_1^2 + \cdots + x_n^2) = d_1^2\,|\mathbf{x}|^2\end{aligned}$$

$\Longrightarrow |\mathbf{Dx}| \,/\, |\mathbf{x}| \le |d_1|$. This shows $\|\mathbf{D}\| \le |d_1|$, i.e., in all $\|\mathbf{D}\| = |d_1|$.

c) Since T is of class C^1, the map $\mathbf{x} \to \mathbf{J}_T(\mathbf{x})$ is continuous. (Its coordinate functions are the partial derivatives $\partial T_i/\partial x_j$.) Since $\mathbf{A} \mapsto \|\mathbf{A}\|$ is continuous as well (the triangle inequality implies $|\|\mathbf{A}\| - \|\mathbf{B}\|| \le \|\mathbf{A} - \mathbf{B}\|$, so that $\delta = \epsilon$ works in a continuity proof), the composition $\mathbf{x} \mapsto \|\mathbf{J}_T(\mathbf{x})\|$ is continuous. Hence, since $\|\mathbf{J}_T(\mathbf{x}^*)\| = \|\mathbf{0}\| = 0$, there exists a ball $\mathrm{B}_r(\mathbf{x}^*)$, $r > 0$, such that $\|\mathbf{J}_T(\mathbf{x})\| < 1/2$ for $\mathbf{x} \in \mathrm{B}_r(\mathbf{x}^*)$. For $\mathbf{x} \in \overline{\mathrm{B}_r(\mathbf{x}^*)}$ we then have $\|\mathbf{J}_T(\mathbf{x})\| \le 1/2$. (The constant $1/2$ is arbitrary; we could achieve $\|\mathbf{J}_T(\mathbf{x})\| < \epsilon$, for any given $\epsilon > 0$, by choosing r suitably.)

As shown in the lecture, we have $T(\mathbf{x}) - T(\mathbf{y}) = \mathbf{A}(\mathbf{x} - \mathbf{y})$ with $\mathbf{A} = \int_0^1 \mathbf{J}_T\big(\mathbf{x} + t(\mathbf{y} - \mathbf{x})\big)\,\mathrm{d}t$. For $\mathbf{x}, \mathbf{y} \in \overline{\mathrm{B}_r(\mathbf{x}^*)}$ and $t \in [0,1]$ we also have $\mathbf{x} + t(\mathbf{y} - \mathbf{x}) \in \overline{\mathrm{B}_r(\mathbf{x}^*)}$ and hence $\big\|\mathbf{J}_T\big(\mathbf{x} + t(\mathbf{y} - \mathbf{x})\big)\big\| \le 1/2$. Since for (continuous) matrix-valued functions $[a,b] \to \mathbb{R}^{n\times n}$, $t \mapsto \mathbf{M}(t)$ the inequality $\left\|\int_a^b \mathbf{M}(t)\,\mathrm{d}t\right\| \le \int_a^b \|\mathbf{M}(t)\|\,\mathrm{d}t$ holds (if you don't believe this, use the corresponding inequality for the Frobenius norm instead and adapt the constants suitably), we obtain $\|\mathbf{A}\| \le \int_0^1 1/2\,\mathrm{d}t = 1/2$ and $|T(\mathbf{x}) - T(\mathbf{y})| \le \frac{1}{2}\,|\mathbf{x} - \mathbf{y}|$ for $\mathbf{x}, \mathbf{y} \in \overline{\mathrm{B}_r(\mathbf{x}^*)}$.

d) We can take $\mathbf{A} = \begin{pmatrix} 3/4 & 0 \\ 0 & 3/4 \end{pmatrix}$. Since $\|\mathbf{A}\| = 3/4$, we have

$$d(\mathbf{Ax}, \mathbf{Ay}) = |\mathbf{Ax} - \mathbf{Ay}| = |\mathbf{A}(\mathbf{x} - \mathbf{y})| \le \|\mathbf{A}\|\,|\mathbf{x} - \mathbf{y}| = \tfrac{3}{4}\,|\mathbf{x} - \mathbf{y}| = \tfrac{3}{4}\,d(\mathbf{x}, \mathbf{y}),$$

so that $\mathbf{x} \mapsto \mathbf{Ax}$ is a contraction. But $\|\mathbf{A}\|_\mathrm{F} = \sqrt{(3/4)^2 + (3/4)^2} = \sqrt{18/16} > 1$.

29 [Note: In the following, the general solutions of the ODE's $y'' \pm y = 0$ are determined using the machinery for higher-order linear ODE's with constant coefficients. You will learn this stuff during the next few weeks. Adhoc derivations of these solutions were given in `lecture1-3_handout.pdf`, Slides 35 f and Exercise H18 a).]
When $y \ge 0$, $y'' + y = 0$. The characteristic polynomial is $X^2 + 1 = 0$ with roots $\lambda_1 = i, \lambda_2 = -i$.
The general real solution is $y(t) = c_1 \cos t + c_2 \sin t$, $c_1, c_2 \in \mathbb{R}$.

$$\because \begin{cases} y(0) = c_1 = 0 \\ y'(0) = c_2 = 1 \end{cases} \quad \therefore y = \sin t \quad (t \in [0, \pi])$$

When $y \le 0$, $y'' - y = 0$. The characteristic polynomial is $X^2 - 1 = 0$ with roots $\lambda_1 = 1, \lambda_2 = -1$.
$\Longrightarrow$ The general real solution is $y(t) = c_1e^t + c_2e^{-t}$, $c_1, c_2 \in \mathbb{R}$.

$$\because \begin{cases} y(0) = c_1 + c_2 = 0 \\ y'(0) = c_1 - c_2 = 1 \end{cases} \Longrightarrow \begin{cases} c_1 = \frac{1}{2} \\ c_2 = -\frac{1}{2} \end{cases} \quad \therefore y(t) = \frac{1}{2}e^t - \frac{1}{2}e^{-t} \quad (t \le 0)$$

In order to get maximal domain $\mathbb{R}$, we impose for $t \ge \pi$ the new initial conditions $y(\pi) = \sin(\pi) = 0$, $y'(\pi) = \cos(\pi) = -1$, which are satisfied by the already defined solution on $(-\infty, \pi]$. Since $y(t) < 0$ for $t \downarrow \pi$, we must fit the general solution for $y \leq 0$, viz. $y(t) = c_1 e^t + c_2 e^{-t}$, to the new initial conditions.

$$\because \begin{cases} y(\pi) = c_1 e^{\pi} + c_2 e^{-\pi} = 0 \\ y'(\pi) = c_1 e^{\pi} - c_2 e^{-\pi} = -1 \end{cases} \Rightarrow \begin{cases} c_1 = -\frac{1}{2e^{\pi}} \\ c_2 = \frac{1}{2e^{-\pi}} \end{cases} \quad \therefore y(t) = -\frac{1}{2e^{\pi}}\,\mathrm{e}^t + \frac{1}{2e^{-\pi}}\,e^{-t} \quad (t \geq \pi)$$

Since this function is negative for all $t > \pi$, it also provides a solution of $y'' + |y| = 0$ on $[\pi, +\infty)$.
The final solution is

$$y(t) = \begin{cases} \frac{1}{2}e^t - \frac{1}{2}e^{-t} = \sinh t & \text{for } t \leq 0, \\ \sin t & \text{for } 0 \leq t \leq \pi, \\ -\frac{1}{2}\,\mathrm{e}^{t-\pi} + \frac{1}{2}\,e^{-(t-\pi)} = -\sinh(t-\pi) & \text{for } t \geq \pi. \end{cases}$$

The function $y(t)$ is differentiable also at $t = 0, \pi$, because the one-sided derivatives exist there and coincide. (In fact $y(t)$ is even C^2, but not C^3.)

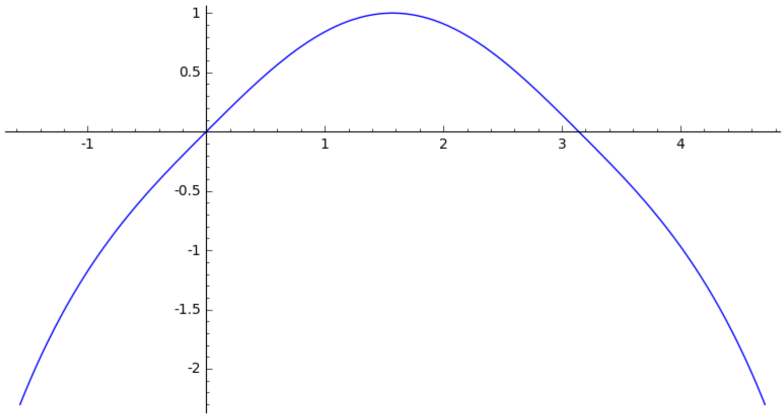

Figure 1: The solution $y(t)$ to H29

The Existence and Uniqueness Theorem applies to $y'' + |y| = 0$, because it is equivalent to the explicit ODE $y'' = f(t, y, y')$ with $f(t, y_0, y_1) = -|y_0|$. The function $f(t, y_0, y_1)$is continuous and satisfies

$$|f(t, y_0, y_1) - f(t, z_0, z_1)| = |-|y_0| + |z_0|| = \pm\big(|y_0| - |z_0|\big) \leq |y_0 - z_0| \leq \sqrt{(y_0 - z_0)^2 + (y_1 - z_1)^2}$$

for all $\mathbf{y} = (y_0, y_1)$, $\mathbf{z} = (z_0, z_1) \in \mathbb{R}^2$, i.e., a global Lipschitz condition with $L = 1$. As shown in the lecture, the (trivially continuous) 1st-order system obtained from $y'' = f(t, y, y')$ by order-reduction then satisfies such a Lipschitz condition as well (perhaps with slightly larger Lipschitz constant), so that the Existence and Uniqueness Theorem can be applied.

30 a) We need to show the missing properties of d required for a metric, i.e., $d(a, b) \geq 0$ and $d(a, b) = d(b, a)$ for all $a, b \in M$. Then we can conclude $d(b, a) = d(a, b) \leq d(b, c)) + d(c, a)$, i.e., the ordinary triangle inequality also holds.

Setting $c = a$ in the postulated "triangle inequality" gives $d(a, b) \leq d(b, a) + d(a, a) = d(b, a)$ for all $a, b \in M$. But then interchanging a and b also yields $d(b, a) \leq d(a, b)$, and so we must have $d(a,b) = d(b,a)$ for all $a, b \in M$. Further, setting $b = a$ in the "triangle inequality" gives $d(a,a) \le d(a,c) + d(c,a)$, which on account of the already proved symmetry of d reduces to $0 \le 2\,d(a,c)$. Thus we also have $d(a,c) \ge 0$ for $a, c \in M$, completing the proof.

b) No. A counterexample is $M = \{a, b\}$ with d defined by $d(a,a) = d(b,b) = 0$, $d(a,b) = 1$, $d(b,a) = -1$. In this case the ordinary triangle inequality has 8 instances:

$$\begin{aligned}
d(a,a) &\le d(a,a) + d(a,a),\\
d(a,a) &\le d(a,b) + d(b,a),\\
d(b,b) &\le d(b,b) + d(b,b),\\
d(b,b) &\le d(b,a) + d(a,b),\\
d(a,b) &\le d(a,a) + d(a,b),\\
d(a,b) &\le d(a,b) + d(b,b),\\
d(b,a) &\le d(b,a) + d(a,a),\\
d(b,a) &\le d(b,b) + d(b,a).
\end{aligned}$$

The only nontrivial relation obtained from these is $d(a,b) + d(b,a) \ge 0$, which is also true in our example. Hence the example satisfies all assumptions made in b).

Remark: This funny exercise is taken from the Book *Set Theory and Metric Spaces* by *Irving Kaplansky*, which is higly recommended for studying if you are interested in the underlying mathematical theory.

31 a) (N1), (N2) follow from the corresponding properties of the Euclidean length. For (N3) this is also true, but here we give a detailed proof: The triangle inequality for $|\cdot|$ yields for $\mathbf{x} \in \mathbb{R}^n$ with $|\mathbf{x}| = 1$ the estimate

$$|(\mathbf{A}+\mathbf{B})\mathbf{x}| = |\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{x}| \le |\mathbf{A}\mathbf{x}| + |\mathbf{B}\mathbf{x}| \le \|\mathbf{A}\| + \|\mathbf{B}\|\,.$$

Taking the maximum over all such vectors $\mathbf{x}$ then gives

$$\|\mathbf{A}+\mathbf{B}\| = \max\bigl\{|(\mathbf{A}+\mathbf{B})\mathbf{x}|\,;\mathbf{x} \in \mathbb{R}^n, |\mathbf{x}| = 1\bigr\} \le \|\mathbf{A}\| + \|\mathbf{B}\|\,.$$

For (N4) we can argue as follows:

$$|(\mathbf{AB})\mathbf{x}| = |\mathbf{A}(\mathbf{B}\mathbf{x})| \le \|\mathbf{A}\|\,|\mathbf{B}\mathbf{x}| \le \|\mathbf{A}\|\,\|\mathbf{B}\|\,|\mathbf{x}| \implies \frac{|(\mathbf{AB})\mathbf{x}|}{|\mathbf{x}|} \le \|\mathbf{A}\|\,\|\mathbf{B}\| \text{ for } \mathbf{x} \ne \mathbf{0}$$

Taking the maximum over all vectors $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ then gives $\|\mathbf{AB}\| \le \|\mathbf{A}\|\,\|\mathbf{B}\|$. (Alternatively we can restrict the above computation to vectors $\mathbf{x}$ of length 1, resulting in $|(\mathbf{AB})\mathbf{x}| \le \|\mathbf{A}\|\,\|\mathbf{B}\|$, and then take the maximum over those vectors as in the proof of (N3).)

b) Since the Frobenius norm is a matrix analogue of the Euclidean length on $\mathbb{R}^{n^2}$, it clearly satisfies (N1)–(N3). For the proof of (N4) we write $\mathbf{A} = (a_{ij})$, $\mathbf{B} = (b_{ij})$, so that $\mathbf{AB} = (c_{ij}) = \bigl(\sum_{k=1}^n a_{ik}b_{kj}\bigr)_{i,j=1}^n$. Denoting the i-th row of $\mathbf{A}$ by $\mathbf{a}_i$ and the j-th column of $\mathbf{B}$ by $\mathbf{b}_j$, we have

$$\begin{aligned}
c_{ij} &= \mathbf{a}_i \cdot \mathbf{b}_j,\\
c_{ij}^2 &\le |\mathbf{a}_i|^2\,|\mathbf{b}_j|^2\,. && \text{(Cauchy-Schwarz Inequality)}
\end{aligned}$$

Summing these inequalities over i, j gives

$$\|\mathbf{AB}\|_{\mathrm{F}}^2 \leq \sum_{i,j=1}^{n} |\mathbf{a}_i|^2 \, |\mathbf{b}_j|^2 = \left(\sum_{i=1}^{n} |\mathbf{a}_i|^2\right)\left(\sum_{j=1}^{n} |\mathbf{b}_j|^2\right) = \|\mathbf{A}\|_{\mathrm{F}}^2 \, \|\mathbf{B}\|_{\mathrm{F}}^2 \,,$$

and (N4) follows.

c) It suffices to show $|\mathbf{Ax}| \leq \|\mathbf{A}\|_{\mathrm{F}}$ for all vectors $\mathbf{x} \in \mathbb{R}^n$ with $|\mathbf{x}| = 1$. Using the notation introduced in b) we have

$$\begin{gathered} \mathbf{Ax} = \begin{pmatrix} \mathbf{a}_1 \cdot \mathbf{x} \\ \vdots \\ \mathbf{a}_n \cdot \mathbf{x} \end{pmatrix}, \\ (\mathbf{a}_i \cdot \mathbf{x})^2 \leq |\mathbf{a}_i|^2 \, |\mathbf{x}|^2 = |\mathbf{a}_i|^2 \,. \\ \implies |\mathbf{Ax}|^2 = \sum_{i=1}^{n} (\mathbf{a}_i \cdot \mathbf{x})^2 \leq \sum_{i=1}^{n} |\mathbf{a}_i|^2 = \|\mathbf{A}\|_{\mathrm{F}}^2 \end{gathered}$$

This proves $|\mathbf{Ax}| \leq \|\mathbf{A}\|_{\mathrm{F}}$ and implies the desired inequality $\|\mathbf{A}\| \leq \|\mathbf{A}\|_{\mathrm{F}}$ for $\mathbf{A} \in \mathbb{R}^{n \times n}$. *Remark:* The matrix norms considered so far and their properties remain true if $\mathbb{R}^n$ is replaced by $\mathbb{C}^n$ and the Euclidean length on $\mathbb{R}^n$ by $|\mathbf{x}| = \sqrt{\sum_{i=1}^{n} |x_i|^2}$. The above proofs remain valid for $\mathbb{C}^n$, provided we change squares of real numbers to squared absolute values of complex numbers, e.g., $(\mathbf{a}_i \cdot \mathbf{x})^2$ becomes $|\mathbf{a}_i \cdot \mathbf{x}|^2$.

d) As shown at the end of `lecture14-16_handout.pdf`, $\|\mathbf{A}\|^2$ is the largest eigenvalue of $\mathbf{A}^{\mathsf{T}}\mathbf{A}$. Applying this to $\mathbf{A}^{\mathsf{T}}$, we see that $\|\mathbf{A}^{\mathsf{T}}\|^2$ is the largest eigenvalue of $\mathbf{A}\mathbf{A}^{\mathsf{T}}$. But the eigenvalues of $\mathbf{A}^{\mathsf{T}}\mathbf{A}$ and $\mathbf{A}\mathbf{A}^{\mathsf{T}}$ are the same (cf. Math257 in Fall 2022, Exercise H46 c) of Homework 9), and hence the same is true of the (spectral) norms of $\mathbf{A}$ and $\mathbf{A}^{\mathsf{T}}$.

e) If $\mathbf{A}$ is invertible then so is $\mathbf{A}^{\mathsf{T}}\mathbf{A}$ (cf. Math257 in Fall 2022, Exercise W26 a) of Worksheet 7). Hence hence all its eigenvalues are positive and $1/\sqrt{\lambda_n}$ is well-defined.

Since $(\mathbf{A}^{-1})^{\mathsf{T}} \mathbf{A}^{-1} = (\mathbf{A}^{\mathsf{T}})^{-1} \mathbf{A}^{-1} = (\mathbf{A}\mathbf{A}^{\mathsf{T}})^{-1}$, the eigenvalues of $(\mathbf{A}^{-1})^{\mathsf{T}} \mathbf{A}^{-1}$ are $\lambda_n^{-1} > \lambda_{n-1}^{-1} > \cdots > \lambda_1^{-1}$ (cf. Math257 in Fall 2022, Exercise W34 a) of Worksheet 10). Thus we have $\|\mathbf{A}^{-1}\|^2 = \lambda_n^{-1}$, as claimed.

f) The entries of $\mathbf{A}^{\mathsf{T}}\mathbf{A}$ are the pairwise dot products of the columns $\mathbf{c}_1, \dots, \mathbf{c}_n$ of $\mathbf{A}$. In particular we have

$$\operatorname{tr}\left(\mathbf{A}^{\mathsf{T}}\mathbf{A}\right) = \sum_{i=1}^{n} \left(\mathbf{A}^{\mathsf{T}}\mathbf{A}\right)_{ii} = |\mathbf{c}_1|^2 + \cdots + |\mathbf{c}_n|^2 = \sum_{i,j=1}^{n} a_{ij}^2 = \|\mathbf{A}\|_{\mathrm{F}}^2 \,.$$

On the other hand, $\operatorname{tr}\left(\mathbf{A}^{\mathsf{T}}\mathbf{A}\right)$ is equal to the sum of the eigenvalues of $\mathbf{A}^{\mathsf{T}}\mathbf{A}$. Both identities taken together give $\|\mathbf{A}\|_{\mathrm{F}}^2 = \lambda_1 + \cdots + \lambda_n$, i.e., $\|\mathbf{A}\|_{\mathrm{F}} = \sqrt{\lambda_1 + \cdots + \lambda_n}$.