MATH 20700 (Honors Analysis in $\mathbb{R}^n$ I) Problem Sets

Steven Labalme

December 2, 2021

Contents

1 Matrix Basics and Linear Systems **1**

2 Eigenvalues and Eigenvectors **14**

3 Inner Product Spaces **23**

4 Inner Product Phenomena and Intro to Bilinear Forms **38**

5 Definiteness, Dual Spaces, and Advanced Spectral Theory **53**

6 Basic Topology **60**

7 Basic Topology II **66**

8 Continuity **75**

9 Sequences and Series of Functions **89**

References **93**

1 Matrix Basics and Linear Systems

From Treil (2017).

Chapter 1

10/4: **1.2.** Which of the following sets (with natural addition and multiplication by a scalar) are vector spaces? Justify your answer.

a) The set of all continuous functions on the interval $[0, 1]$.

Answer. This IS a vector space.

Commutativity: If f, g are continuous functions on $[0, 1]$, then $f + g$ is continuous on $[0, 1]$ with $f + g = g + f$.

Associativity: If f, g, h are continuous functions on $[0, 1]$, then $(f + g) + h$ and $f + (g + h)$ are continuous functions on $[0, 1]$ with $(f + g) + h = f + (g + h)$.

Zero vector: Let $\mathbf{0} : [0, 1] \to [0, 1]$ be defined by $\mathbf{0}(x) = 0$ for all $x \in [0, 1]$. Then if f is any continuous function on $[0, 1]$, $f + \mathbf{0} = f$.

Additive inverse: Let f be a continuous function on $[0, 1]$. Define $g : [0, 1] \to [0, 1]$ by $g(x) = -f(x)$ for all $x \in [0, 1]$. Clearly g is still continuous on $[0, 1]$, and $f + g = \mathbf{0}$.

Multiplicative identity: Let f be a continuous function on $[0, 1]$. Then naturally $1f = f$.

Multiplicative associativity: Let f be a continuous function on $[0, 1]$, and let $\alpha, \beta \in \mathbb{F}$. Then $(\alpha\beta)f = \alpha(\beta f)$.

Distributive (vectors): Let f, g be continuous on $[0, 1]$, and let $\alpha \in \mathbb{F}$. Then $\alpha(f + g)$ and $\alpha f + \alpha g$ are continuous on $[0, 1]$ and equal.

Distributive (scalars): Let f be continuous on $[0, 1]$, and let $\alpha, \beta \in \mathbb{F}$. Then $(\alpha + \beta)f$ and $\alpha f + \beta f$ are continuous on $[0, 1]$ and equal. □

b) The set of all non-negative functions on the interval $[0, 1]$.

Answer. This IS NOT a vector space.

Not closed under inverses — $f : [0, 1] \to \mathbb{R}$ defined by $f(x) = 1$ for all x would be a non-negative function on this interval, and $g : [0, 1] \to \mathbb{R}$ defined by $g(x) = -1$ for all x is naturally its inverse, but not an element of the set. □

c) The set of all polynomials of degree *exactly* n.

Answer. This IS NOT a vector space.

Not closed under summation — the inverse of x^n is $-x^n$, but their sum is 0, a polynomial of degree 0. □

d) The set of all symmetric $n \times n$ matrices, i.e., the set of matrices $A = \{a_{j,k}\}_{j,k=1}^{n}$ such that $A^T = A$.

Answer. This IS a vector space.

The condition for symmetric is $a_{j,k} = a_{k,j}$. Assume this is true for A and B. Then naturally

$$\begin{aligned}(a + b)_{j,k} &= a_{j,k} + b_{j,k} \\ &= a_{k,j} + b_{k,j} \\ &= (a + b)_{k,j}\end{aligned}$$

A symmetric argument verifies scalar multiplication. □

1.3. True or false:

a) Every vector space contains a zero vector.

Answer. True.
By definition. □

b) A vector space can have more than one zero vector.

Answer. False.
Suppose for the sake of contradiction that $0, 0'$ are two distinct zero vectors. Then

$$0 = 0 + 0' = 0'$$

a contradiction. □

c) An $m \times n$ matrix has m rows and n columns.

Answer. True.
By definition. □

d) If f and g are polynomials of degree n, then $f + g$ is also a polynomial of degree n.

Answer. False.
x^n and $-x^n$ are both polynomials of degree n, but their sum (0) is a polynomial of degree 0. □

e) If f and g are polynomials of degree at most n, then $f + g$ is also a polynomial of degree at most n.

Answer. True.
Suppose for the sake of contradiction that there exist f, g of degree at most n such that $f + g$ has degree $m > n$. Then $f + g$ has an ax^m term. Since f has degree n, it has no bx^m term. Thus, $(f + g) - f = g$ retains the ax^m term, and is of degree $m > n$, a contradiction. □

2.2. True or false:

a) Any set containing a zero vector is linearly dependent.

Answer. True.
Let $\mathbf{v}_1, \dots, \mathbf{v}_n$ be a list of vectors. If $\mathbf{v}_i = 0$, then

$$0\mathbf{v}_1 + \cdots + 0\mathbf{v}_{i-1} + 1\mathbf{v}_i + 0\mathbf{v}_{i+1} + \cdots + 0\mathbf{v}_n = \mathbf{0}$$

even though one of the coefficients isn't 0. Thus, the list is linearly dependent. □

b) A basis must contain $\mathbf{0}$.

Answer. False.
$\{1\}$ is a basis of $\mathbb{R}^1$. □

c) Subsets of linearly dependent sets are linearly dependent.

Proof. False.
$\{\left[\begin{smallmatrix}0\\0\end{smallmatrix}\right], \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]\}$ is linearly dependent, but $\{\left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]\}$ is linearly independent. □

d) Subsets of linearly independent sets are linearly independent.

Proof. True.
Suppose for the sake of contradiction that there exists a linearly dependent subset of a linearly independent list. Then there are nonzero coefficients that make a linear combination of the linearly dependent equal to zero. Thus, if we pair these coefficients to their respective vectors in a sum of the whole list, and use zero everywhere else, we will have a set of coefficients, not all zero, that make the supposedly linearly independent list sum to zero, a contradiction. □

e) If $\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n = \mathbf{0}$, then all scalars α_k are zero.

Answer. False.
Let $\mathbf{v}_1, \mathbf{v}_2$ be defined by

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \mathbf{v}_2 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

Then $1\mathbf{v}_1 + 1\mathbf{v}_2 = \mathbf{0}$. □

2.5. Let a system of vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ be linearly independent but not generating. Show that it is possible to find a vector $\mathbf{v}_{r+1}$ such that the system $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r, \mathbf{v}_{r+1}$ is linearly independent. (Hint: Take for $\mathbf{v}_{r+1}$ any vector that cannot be represented as a linear combination $\sum_{k=1}^{r} \alpha_k\mathbf{v}_k$ and show that the system $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r, \mathbf{v}_{r+1}$ is linearly independent.)

Answer. Let $\mathbf{v}_{r+1}$ be any vector that cannot be represented as a linear combination $\sum_{k=1}^{r} \alpha_k\mathbf{v}_k$ (we are guaranteed that one exists, because otherwise $\mathbf{v}_1, \ldots, \mathbf{v}_r$ would be generating). Now suppose for the sake of contradiction that the new list is linearly dependent. Then there exist coefficients $\alpha_1, \ldots, \alpha_{r+1}$, not all zero, such that

$$\alpha_1\mathbf{v}_1 + \cdots + \alpha_{r+1}\mathbf{v}_{r+1} = \mathbf{0}$$

We divide into two cases ($\alpha_{r+1} = 0$ and $\alpha_{r+1} \neq 0$). If $\alpha_{r+1} = 0$, then it must be at least one of $\alpha_1, \ldots, \alpha_r$ is nonzero by hypothesis. But then

$$\begin{aligned} \mathbf{0} &= \alpha_1\mathbf{v}_1 + \cdots + \alpha_r\mathbf{v}_r + \alpha_{r+1}\mathbf{v}_{r+1} \\ \mathbf{0} - 0\mathbf{v}_{r+1} &= \alpha_1\mathbf{v}_1 + \cdots + \alpha_r\mathbf{v}_r \\ \mathbf{0} &= \alpha_1\mathbf{v}_1 + \cdots + \alpha_r\mathbf{v}_r \end{aligned}$$

for a set of coefficients $\alpha_1, \ldots, \alpha_r \in \mathbb{F}$, not all zero, contradicting the fact that $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are linearly independent. On the other hand, if $\alpha_{r+1} \neq 0$, then

$$\mathbf{v}_{r+1} = -\frac{\alpha_1}{\alpha_{r+1}}\mathbf{v}_1 - \cdots - \frac{\alpha_r}{\alpha_{r+1}}\mathbf{v}_r$$

so $\mathbf{v}_{r+1}$ can be expressed as a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_r$, a contradiction. □

2.6. Is it possible that vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are linearly dependent, but the vectors $\mathbf{w}_1 = \mathbf{v}_1 + \mathbf{v}_2, \mathbf{w}_2 = \mathbf{v}_2 + \mathbf{v}_3, \mathbf{w}_3 = \mathbf{v}_3 + \mathbf{v}_1$ are linearly *independent*?

Answer. No.

Suppose $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are linearly dependent. Then there exist coefficients $\alpha_1, \alpha_2, \alpha_3$, not all zero, such that

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3 = \mathbf{0}$$

To prove that $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$ must be linearly dependent, as defined, it will suffice to show that there exist coefficients $\beta_1, \beta_2, \beta_3$, not all zero, such that

$$\beta_1\mathbf{w}_1 + \beta_2\mathbf{w}_2 + \beta_3\mathbf{w}_3 = \mathbf{0}$$

But we have that

$$\begin{aligned} \beta_1\mathbf{w}_1 + \beta_2\mathbf{w}_2 + \beta_3\mathbf{w}_3 &= \beta_1(\mathbf{v}_1 + \mathbf{v}_2) + \beta_2(\mathbf{v}_2 + \mathbf{v}_3) + \beta_3(\mathbf{v}_3 + \mathbf{v}_1) \\ &= (\beta_1 + \beta_3)\mathbf{v}_1 + (\beta_1 + \beta_2)\mathbf{v}_2 + (\beta_2 + \beta_3)\mathbf{v}_3 \end{aligned}$$

so to have $(\beta_1 + \beta_3)\mathbf{v}_1 + (\beta_1 + \beta_2)\mathbf{v}_2 + (\beta_2 + \beta_3)\mathbf{v}_3 = \mathbf{0}$, we need only require that

$$\beta_1 + \beta_3 = \alpha_1 \qquad \beta_1 + \beta_2 = \alpha_2 \qquad \beta_2 + \beta_3 = \alpha_3$$

Thus, choose

$$\beta_1 = \frac{1}{2}(\alpha_1 + \alpha_2 - \alpha_3) \qquad \beta_2 = \frac{1}{2}(-\alpha_1 + \alpha_2 + \alpha_3) \qquad \beta_3 = \frac{1}{2}(\alpha_1 - \alpha_2 + \alpha_3)$$

Lastly, note that we do not have $\beta_1 = \beta_2 = \beta_3 = 0$ because if we did, we could prove from that condition that $\alpha_1 = \alpha_2 = \alpha_3 = 0$, a contradiction. $\square$

3.3. For each linear transformation below, find its matrix.

a) $T : \mathbb{R}^2 \to \mathbb{R}^3$ defined by $T(x, y)^T = (x + 2y, 2x - 5y, 7y)^T$.

Answer.

$$\begin{pmatrix} 1 & 2 \\ 2 & -5 \\ 0 & 7 \end{pmatrix}$$

$\square$

b) $T : \mathbb{R}^4 \to \mathbb{R}^3$ defined by $T(x_1, x_2, x_3, x_4)^T = (x_1 + x_2 + x_3 + x_4, x_2 - x_4, x_1 + 3x_2 + 6x_4)^T$.

Answer.

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & -1 \\ 1 & 3 & 0 & 6 \end{pmatrix}$$

$\square$

c) $T : \mathbb{P}_n \to \mathbb{P}_n$ defined by $Tf(t) = f'(t)$ (find the matrix with respect to the standard basis $1, t, t^2, \dots, t^n$).

Answer.

$$\begin{pmatrix} 0 & 1 & 0 & & 0 \\ 0 & 0 & 2 & & 0 \\ & & & \ddots & \\ 0 & 0 & 0 & & n \\ 0 & 0 & 0 & & 0 \end{pmatrix}$$

$\square$

d) $T : \mathbb{P}_n \to \mathbb{P}_n$ defined by $Tf(t) = 2f(t) + 3f'(t) - 4f''(t)$ (again with respect to the standard basis $1, t, t^2, \dots, t^n$).

Answer.

$$\begin{pmatrix} 2 & 3 & -8 & & 0 \\ 0 & 2 & 6 & \ddots & 0 \\ 0 & 0 & 2 & \ddots & -4n(n-1) \\ \vdots & \vdots & & \ddots & 3n \\ 0 & 0 & 0 & & 2 \end{pmatrix}$$

$\square$

3.6. The set $\mathbb{C}$ of complex numbers can be canonically identified with the space $\mathbb{R}^2$ by treating each $z = x + iy \in \mathbb{C}$ as a column $(x, y)^T \in \mathbb{R}^2$.

a) Treating $\mathbb{C}$ as a complex vector space, show that the multiplication by $\alpha = a + ib \in \mathbb{C}$ is a linear transformation in $\mathbb{C}$. What is its matrix?

Answer. Let $T : \mathbb{C} \to \mathbb{C}$ be defined by $Tx = \alpha x$. Then

$$\begin{aligned} T(x+y) &= \alpha(x+y) \\ &= \alpha x + \alpha y \\ &= Tx + Ty \end{aligned} \qquad \begin{aligned} T(\beta x) &= \alpha(\beta x) \\ &= \beta(\alpha x) \\ &= \beta Tx \end{aligned}$$

so T is linear. The matrix of T is $[\alpha]$. □

b) Treating $\mathbb{C}$ as the real vector space $\mathbb{R}^2$, show that the multiplication by $\alpha = a + ib$ defines a linear transformation there. What is its matrix?

Answer. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $T(x,y)^T = (ax - by, ay + bx)^T$. Then

$$\begin{aligned} T\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) &= T\left(\begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \end{pmatrix}\right) \\ &= \begin{pmatrix} a(x_1+y_1) - b(x_2+y_2) \\ a(x_2+y_2) + b(x_1+y_1) \end{pmatrix} \\ &= \begin{pmatrix} ax_1 - bx_2 \\ ax_2 + bx_1 \end{pmatrix} + \begin{pmatrix} ay_1 - by_2 \\ ay_2 + by_1 \end{pmatrix} \\ &= T\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + T\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \end{aligned} \qquad \begin{aligned} T\left(c\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) &= T\left(\begin{pmatrix} cx_1 \\ cx_2 \end{pmatrix}\right) \\ &= \begin{pmatrix} a(cx_1) - b(cx_2) \\ a(cx_2) + b(cx_1) \end{pmatrix} \\ &= c\begin{pmatrix} ax_1 - bx_2 \\ ax_2 + bx_1 \end{pmatrix} \\ &= cT\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \end{aligned}$$

so T is linear. The matrix of T is

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

□

c) Define $T(x+iy) = 2x - y + i(x - 3y)$. Show that this transformation is not a linear transformation in the complex vector space $\mathbb{C}$, but if we treat $\mathbb{C}$ as the real vector space $\mathbb{R}^2$, then it is a linear transformation there (i.e., that T is a *real linear* but not a *complex linear* transformation). Find the matrix of the real linear transformation T.

Answer. To prove that T is not complex linear, note that

$$T(i \cdot 1) = T(i) = -1 - 3i \neq -1 + 2i = i(2+i) = iT(1)$$

We can verify the T is real linear with the following.

$$\begin{aligned} T\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) &= T\begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \end{pmatrix} \\ &= \begin{pmatrix} 2(x_1+y_1) - (x_2+y_2) \\ (x_1+y_1) - 3(x_2+y_2) \end{pmatrix} \\ &= \begin{pmatrix} 2x_1 - x_2 \\ x_1 - 3x_2 \end{pmatrix} + \begin{pmatrix} 2y_1 - y_2 \\ y_1 - 3y_2 \end{pmatrix} \\ &= T\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + T\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \end{aligned} \qquad \begin{aligned} T\left(c\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) &= T\begin{pmatrix} cx_1 \\ cx_2 \end{pmatrix} \\ &= \begin{pmatrix} 2(cx_1) - (cx_2) \\ (cx_1) - 3(cx_2) \end{pmatrix} \\ &= c\begin{pmatrix} 2x_1 - x_2 \\ x_1 - 3x_2 \end{pmatrix} \\ &= cT\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \end{aligned}$$

The matrix of the real linear transformation is the following.

$$\begin{pmatrix} 2 & -1 \\ 1 & -3 \end{pmatrix}$$

□

5.3. Multiply two rotation matrices T_α and T_β (it is a rare case when the multiplication is commutative, i.e., $T_\alpha T_\beta = T_\beta T_\alpha$, so the order is not essential). Deduce formulas for $\sin(\alpha+\beta)$ and $\cos(\alpha+\beta)$ from here.

Answer.

$$\begin{aligned}
T_\alpha T_\beta &= \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{pmatrix} \\
&= \begin{pmatrix} \cos\alpha\cos\beta - \sin\alpha\sin\beta & -\cos\alpha\sin\beta - \sin\alpha\cos\beta \\ \sin\alpha\cos\beta + \cos\alpha\sin\beta & -\sin\alpha\sin\beta + \cos\alpha\cos\beta \end{pmatrix}
\end{aligned}$$

Since $T_{\alpha+\beta} = T_\alpha T_\beta$, we have that

$$\begin{aligned}
\sin(\alpha+\beta) &= \sin\alpha\cos\beta + \cos\alpha\sin\beta \\
\cos(\alpha+\beta) &= \cos\alpha\cos\beta - \sin\alpha\sin\beta
\end{aligned}$$

□

5.5. Find linear transformations $A, B : \mathbb{R}^2 \to \mathbb{R}^2$ such that $AB = \mathbf{0}$ but $BA \neq \mathbf{0}$.

Answer. Let

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix}$$

Then

$$AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \qquad BA = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$$

□

5.8. Find the matrix of the reflection through the line $y = -2x/3$. Perform all the multiplications.

Answer. The reflection matrix T can be obtained by composing a rotation of $\mathbb{R}^2$ such that $y = -2x/3$ lines up with the x-axis, a reflection over the x-axis (a super simple reflection), and a rotation back. Let γ be the angle between the x-axis and the line $y = -2x/3$. Then

$$\begin{aligned}
T &= R_{-\gamma} T_0 R_\gamma \\
&= \begin{pmatrix} \cos(-\gamma) & -\sin(-\gamma) \\ \sin(-\gamma) & \cos(-\gamma) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \cos\gamma & -\sin\gamma \\ \sin\gamma & \cos\gamma \end{pmatrix} \\
&= \begin{pmatrix} \cos\gamma & \sin\gamma \\ -\sin\gamma & \cos\gamma \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \cos\gamma & -\sin\gamma \\ \sin\gamma & \cos\gamma \end{pmatrix} \\
&= \begin{pmatrix} \frac{3}{\sqrt{13}} & \frac{2}{\sqrt{13}} \\ -\frac{2}{\sqrt{13}} & \frac{3}{\sqrt{13}} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \frac{3}{\sqrt{13}} & -\frac{2}{\sqrt{13}} \\ \frac{2}{\sqrt{13}} & \frac{3}{\sqrt{13}} \end{pmatrix} \\
&= \begin{pmatrix} \frac{3}{\sqrt{13}} & \frac{2}{\sqrt{13}} \\ -\frac{2}{\sqrt{13}} & \frac{3}{\sqrt{13}} \end{pmatrix} \begin{pmatrix} \frac{3}{\sqrt{13}} & -\frac{2}{\sqrt{13}} \\ -\frac{2}{\sqrt{13}} & -\frac{3}{\sqrt{13}} \end{pmatrix} \\
&= \begin{pmatrix} \frac{5}{13} & -\frac{12}{13} \\ -\frac{12}{13} & -\frac{5}{13} \end{pmatrix}
\end{aligned}$$

□

6.3. Find all left inverses of the column $(1, 2, 3)^T$.

Answer. The set of all left inverses of $(1,2,3)^T$ is the set of all 1×3 matrices (a,b,c) such that $(a,b,c)\cdot(1,2,3)^T = (1)$. In other words, it's the set of all (a,b,c) such that $a+2b+3c=1$. □

6.6. Suppose the product AB is invertible. Show that A is right invertible and B is left invertible. (Hint: You can just write formulas for right and left inverses.)

Answer. If AB is invertible, then there exists $(AB)^{-1}$. It follows that $(AB)(AB)^{-1} = A(B(AB)^{-1}) = I$, so A is right invertible, and $(AB)^{-1}(AB) = ((AB)^{-1}A)B = I$, so B is left invertible. □

6.8. Let A be an $n\times n$ matrix. Prove that if $A^2 = \mathbf{0}$, then A is not invertible.

Answer. Suppose for the sake of contradiction there exists an A^{-1}. Then

$$I = AAA^{-1}A^{-1} = A^2A^{-2} = \mathbf{0}A^{-2} = \mathbf{0}$$

a contradiction. □

6.10. Write matrices of the linear transformations T_1 and T_2 in $\mathbb{F}^5$, defined as follows: T_1 interchanges the coordinates x_2 and x_4 of the vector $\mathbf{x}$, and T_2 just adds to the coordinate x_2 the quantity a times the coordinate x_4, and does not change other coordinates, i.e.,

$$T_1\begin{pmatrix}x_1\\x_2\\x_3\\x_4\\x_5\end{pmatrix} = \begin{pmatrix}x_1\\x_4\\x_3\\x_2\\x_5\end{pmatrix} \qquad T_2\begin{pmatrix}x_1\\x_2\\x_3\\x_4\\x_5\end{pmatrix} = \begin{pmatrix}x_1\\x_2+ax_4\\x_3\\x_4\\x_5\end{pmatrix}$$

where a is some fixed number. Show that T_1 and T_2 are invertible transformations, and write the matrices of the inverses. (Hint: It may be simpler, if you first describe the inverse transformation, and then find its matrix, rather than trying to guess [or compute] the inverses of the matrices T_1, T_2.)

Answer.

$$T_1 = \begin{pmatrix}1&0&0&0&0\\0&0&0&1&0\\0&0&1&0&0\\0&1&0&0&0\\0&0&0&0&1\end{pmatrix} \qquad T_2 = \begin{pmatrix}1&0&0&0&0\\0&1&0&a&0\\0&0&1&0&0\\0&0&0&1&0\\0&0&0&0&1\end{pmatrix}$$

The inverse transformation of T_1 exchanges x_2 and x_4 back, leaving everything else the same. The inverse transformation of T_2 subtracts ax_4 from the second slot, leaving everything else the same. Thus,

$$T_1^{-1} = \begin{pmatrix}1&0&0&0&0\\0&0&0&1&0\\0&0&1&0&0\\0&1&0&0&0\\0&0&0&0&1\end{pmatrix} \qquad T_2^{-1} = \begin{pmatrix}1&0&0&0&0\\0&1&0&-a&0\\0&0&1&0&0\\0&0&0&1&0\\0&0&0&0&1\end{pmatrix}$$

□

6.13. Let A be an invertible symmetric ($A^T = A$) matrix. Is the inverse of A symmetric? Justify.

Answer. We have that

$$\begin{aligned}A^{-1} &= ((A^{-1})^T)^T\\ &= ((A^T)^{-1})^T\\ &= (A^{-1})^T\end{aligned}$$

as desired. □

7.3. Let X be a subspace of a vector space V, and let $\mathbf{v} \in V$, $\mathbf{v} \notin X$. Prove that if $\mathbf{x} \in X$, then $\mathbf{x}+\mathbf{v} \notin X$.

Answer. Suppose for the sake of contradiction that $\mathbf{x}+\mathbf{v} \in X$. Then $\mathbf{x}+\mathbf{v}$ can be expressed as a linear combination of a basis of X. Similarly, $\mathbf{x}$ can be expressed as a linear combination of a basis of X. But this implies that $\mathbf{v} = \mathbf{x}+\mathbf{v}-\mathbf{x}$ can be written as a linear combination of the basis of X, a contradiction since $\mathbf{v} \notin X$, so it shouldn't be able to be written as a linear combination of a basis of X. $\square$

7.4. Let X and Y be subspaces of a vector space V. Using the previous exercise, show that $X \cup Y$ is a subspace if and only if $X \subset Y$ or $Y \subset X$.

Answer. Suppose that $X \cup Y$ is a subspace of V. Suppose for the sake of contradiction that $X \not\subset Y$ and $Y \not\subset X$. Then there exists $\mathbf{x} \in X$ such that $\mathbf{x} \notin Y$ and $\mathbf{y} \in Y$ such that $\mathbf{y} \notin X$. Consider $\mathbf{x}+\mathbf{y}$. Since $\mathbf{x} \in X$ and $\mathbf{y} \notin X$, we have by 7.3 that $\mathbf{x}+\mathbf{y} \notin X$. Similarly, we have that $\mathbf{x}+\mathbf{y} \notin Y$. But this implies that $\mathbf{x}+\mathbf{y} \notin X \cup Y$, contradiction the hypothesis that $X \cup Y$ is a subspace (and thus closed under addition).

Suppose that $X \subset Y$. To prove that $X \cup Y$ is a subspace, it will suffice to check that $\mathbf{v} \in X \cup Y$ implies $\alpha\mathbf{v} \in X \cup Y$, and $\mathbf{v}, \mathbf{w} \in X \cup Y$ implies $\mathbf{v}+\mathbf{w} \in X \cup Y$. Let $\mathbf{v} \in X \cup Y$. Then $\mathbf{v} \in X$ or $\mathbf{v} \in Y$. Either way, the fact that X and Y are subspaces guarantees that $\alpha\mathbf{v} \in X \cup Y$. Now let $\mathbf{v}, \mathbf{w} \in X \cup Y$. Since $X \subset Y$, this implies that $\mathbf{v}, \mathbf{w} \in Y$, so $\mathbf{v}+\mathbf{w} \in Y$, so $\mathbf{v}+\mathbf{w} \in X \cup Y$. The proof is symmetric if $Y \subset X$. $\square$

7.5. What is the smallest subspace of the space of 4×4 matrices which contains all upper triangular matrices ($a_{j,k} = 0$ for all $j > k$), and all symmetric matrices ($A = A^T$)? What is the largest subspace contained in both of those subspaces?

Answer. Out of the vector space V of 4×4 matrices, the smallest subspace which contains all upper triangular matrices and all symmetric matrices is V, itself. This is because any matrix can be decomposed into the sum of a symmetric matrix and an upper triangular matrix (fix the values in the lower triangle, and modify the upper triangle as needed with the upper triangular matrix), so every 4×4 matrix is in this subspace.

The largest subspace contained in both the subspace of upper triangular matrices and the subspace of all symmetric matrices is the subspace of all diagonal matrices. Adding another dimension by making a value *below* the diagonal nonzero makes the matrix in question not upper triangular, and adding another dimension by making a value *above* the diagonal nonzero makes the matrix not symmetric (as we would have to add a value below the diagonal to make it so and that would run into the problem described first). $\square$

Chapter 2

3.4. Do the polynomials x^3+2x, x^2+x+1, x^3+5 generate (span) $\mathbb{P}_3$? Justify your answer.

Answer. $1, x, x^2, x^3$ is the standard basis of $\mathbb{P}_3$. Thus, it spans $\mathbb{P}_3$. But since the given list has fewer vectors, Proposition 3.5 asserts that it cannot span $\mathbb{P}_3$. $\square$

3.5. Can 5 vectors in $\mathbb{F}^4$ be linearly independent? Justify your answer.

Answer. No — see Proposition 3.2. $\square$

3.7. Prove or disprove: If the columns of a square ($n \times n$) matrix A are linearly independent, so are the rows of $A^3 = AAA$.

Answer. Suppose A is $n \times n$ with linearly independent columns. Then by Proposition 3.1, A_e has a pivot in every column. But since A_e is square, this means it also has a pivot in every row. It follows by 3.6 that A is invertible. Thus A^{-1} exists. Consequently, A^{-3} is the inverse of A^3 since

$$A^3A^{-3} = AAAA^{-1}A^{-1}A^{-1} = I \qquad\qquad A^{-3}A^3 = A^{-1}A^{-1}A^{-1}AAA = I$$

so A^3 is invertible. Thus, 3.6 implies A^3_e has a pivot in every row and column. But this implies that $(A^3)^T$ has a pivot in every row and column, meaning by 3.1 that the columns of $(A^3)^T$ are linearly independent, i.e., the rows of A^3 are linearly independent. ☐

5.1. True or false:

a) Every vector space that is generated by a finite set has a basis.

Answer. True.
See Proposition 2.8, Chapter 1. ☐

b) Every vector space has a (finite) basis.

Answer. False.
Consider the vector space of polynomials of any degree. ☐

c) A vector space cannot have more than one basis.

Answer. False.
Both 1 and 2 are bases of $\mathbb{R}^1$. ☐

d) If a vector space has a finite basis, then the number of vectors in every basis is the same.

Answer. True.
See Proposition 3.3, Chapter 2 ☐

e) The dimension of $\mathbb{P}_n$ is n.

Answer. False.
The standard basis of $\mathbb{P}_n$ is $1, t, t^2, \dots, t^n$, which has $n+1$ vectors. Thus, $\dim \mathbb{P}_n = n+1$. ☐

f) The dimension on $M_{m\times n}$ is $m+n$.

Answer. False.
The standard basis of $M_{m\times n}$ is the set of all matrices with a 1 in one slot and a 0 everywhere else. Thus, $\dim M_{m\times n} = m \times n$. ☐

g) If vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ generate (span) the vector space V, then every vector in V can be written as a linear combination of vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ in only one way.

Answer. False.
The vectors $1, 2 \in \mathbb{R}^1$ span $\mathbb{R}^1$, but $3 = 1+2$ and $3 = -1(1) + 2(2)$. ☐

h) Every subspace of a finite-dimensional space is finite-dimensional.

Answer. True.
See Theorem 5.5. ☐

i) If V is a vector space having dimension n, then V has exactly one subspace of dimension 0 and exactly one subspace of dimension n.

Answer. True.
$\{0\}$ is THE unique VS of dimension 0 and a subspace of every vector space, so that part is true. On the other hand, any subspace of $\dim n$ has a basis consisting of n linearly independent, spanning elements of V. But any such list is also a basis of V, so the subspace is V. ☐

5.2. Prove that if V is a vector space having dimension n, then a system of vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ in V is linearly independent if and only if it spans V.

Answer. Suppose first that $\mathbf{v}_1, \dots, \mathbf{v}_n$ is linearly independent. Then the $n \times n$ matrix A with these vectors as columns has a pivot in every column by 3.1. But since A is square, this means that it has a pivot in every row. Thus, by 3.1 again, the columns (i.e, the list $\mathbf{v}_1, \dots, \mathbf{v}_n$) spans V.

The proof is the same in the reverse direction. □

5.6. Consider in the space $\mathbb{R}^5$ vectors $\mathbf{v}_1 = (2, -1, 1, 5, -3)^T$, $\mathbf{v}_2 = (3, -2, 0, 0, 0)^T$, $\mathbf{v}_3 = (1, 1, 50, -921, 0)^T$. (Hint: If you do part (b) first, you can do everything without any computations.)

a) Prove that these vectors are linearly independent.

Answer. If we add in $\mathbf{e}_1$ and $\mathbf{e}_3$ to the mix, then we can create the matrix

$$A = \begin{pmatrix} 1 & 3 & 0 & 1 & 2 \\ 0 & -2 & 0 & 1 & -1 \\ 0 & 0 & 1 & 50 & 1 \\ 0 & 0 & 0 & -921 & 5 \\ 0 & 0 & 0 & 0 & -3 \end{pmatrix}$$

A is already in eschelon form ($A = A_e$) and $A = A_e$ has a pivot in every column, so 3.1 implies that the vectors of A are linearly independent. □

b) Complete the system of vectors to a basis.

Answer. Using the same matrix as above, we can see that A has a pivot in every row and column, so 3.1 implies that its columns form a basis. Thus, the two vectors we added complete the system to a basis of $\mathbb{R}^5$. □

6.1. True or false:

a) Any system of linear equations has at least one solution.

Answer. False.
$y = x$ and $y = x + 1$ has no solution. □

b) Any system of linear equations has at most one solution.

Answer. False.
$y = x$ and $y = x$ has infinite solutions. □

c) Any homogeneous system of linear equations has at least one solution.

Answer. True.
$\mathbf{0}$ is always a solution. □

d) Any system of n linear equations in n unknowns has at least one solution.

Answer. False.
$y = x$ and $y = x + 1$ is a system of 2 linear equations in 2 unknowns but has no solution. □

e) Any system of n linear equations in n unknowns has at most one solution.

Answer. False.
$y = x$ and $y = x$ is a system of 2 linear equations in 2 unknowns but has infinite solutions. □

f) If the homogeneous system corresponding to a given system of linear equations has a solution, then the given system has a solution.

Answer. False.
$y = x$ and $y = x$ is the homogeneous system corresponding to $y = x$ and $y = x + 1$, and it has a solution, but the system itself does not. □

g) If the coefficient matrix of a homogeneous system of n linear equations in n unknowns is invertible, then the system has no non-zero solutions.

Answer. True.
Invertible implies pivots in every row/column by 3.1. This implies that A_{re} gives $\mathbf{0}$ as a particular solution, and the only solution to $A\mathbf{x} = \mathbf{b} = \mathbf{0}$. Thus, 6.1 implies that the set of all solutions is $\{\mathbf{x} + \mathbf{y} : \mathbf{x} \in \{\mathbf{0}\}, \mathbf{y} \in \{\mathbf{0}\}\} = \{\mathbf{0}\}$. □

h) The solution set of any system of m equations in n unknowns is a subspace of $\mathbb{R}^n$.

Answer. False.
The system $x + y = 1$ and $2x + y = 1$ has one solution in $\mathbb{R}^2$, namely $\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)$. Since $\mathbf{0} \notin \{\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)\}$, the solution set is not a *subspace* of $\mathbb{R}^2$, a contradiction. □

i) The solution set of any homogeneous system of m equations in n unknowns is a subspace of $\mathbb{R}^n$.

Answer. True.
Let X be the solution set and let A be the coefficient matrix. The answer to Problem 6.1c shows that $\mathbf{0} \in X$. If $\mathbf{x} \in X$ and $\alpha \in \mathbb{F}$, then $A(\alpha\mathbf{x}) = \alpha A\mathbf{x} = \alpha\mathbf{0} = \mathbf{0}$, so $\alpha\mathbf{x} \in X$. If $\mathbf{x}, \mathbf{y} \in X$, then $A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y} = \mathbf{0} + \mathbf{0} = \mathbf{0}$, so $\mathbf{x} + \mathbf{y} \in X$. □

7.1. True or false:

a) The rank of a matrix is equal to the number of its non-zero columns.

Answer. False.
The rank of a matrix is equal to the number of its pivot columns since each pivot column of A is a vector in the basis of $\operatorname{Ran} A$. In particular, note that non-zero columns can still be linearly dependent. □

b) The $m \times n$ zero matrix is the only $m \times n$ matrix having rank 0.

Answer. True.
Suppose for the sake of contradiction that there exists a nonzero matrix with rank 0. The first column from the left with a nonzero entry will be a pivot column. Thus, this column will be part of the basis of $\operatorname{Ran} A$. But since this column exists, $\operatorname{Ran} A \geq 1$, a contradiction. □

c) Elementary row operations preserve rank.

Answer. True.
Elementary row operations, as left multiplictions by invertible matrices, do not affect linear independence. □

d) Elementary column operations do not necessarily preserve rank.

Answer. False.
Elementary column operations are the same as elementary row operations on the transpose, which we know preserve rank by the above. □

e) The rank of a matrix is equal to the maximum number of linearly independent columns in the matrix.

Answer. True.
Each pivot column is linearly independent, and the rank is equal to the number of pivot columns. □

f) The rank of a matrix is equal to the maximum number of linearly independent rows in the matrix.

Answer. True.
Each pivot row is linearly independent, and the rank is equal to the number of pivot rows/columns. □

g) The rank of an $n \times n$ matrix is at most n.

Answer. True.
Each linearly independent column contributes $+1$ to the rank, and since an $n \times n$ matrix can have at most n *columns*, it certainly cannot have more than n *linearly independent columns*. □

h) An $n \times n$ matrix having rank n is invertible.

Answer. True.
If an $n \times n$ matrix has rank n, then it has n pivot columns. But this implies by 3.6 that it is invertible. □

7.4. Prove that if $A : X \to Y$ and V is a subspace of X, then $\dim AV \leq \operatorname{rank} A$. ($AV$ here means the subspace V transformed by the transformation A, i.e., any vector in AV can be represented as $A\mathbf{v}$, $\mathbf{v} \in V$.) Deduce from here that $\operatorname{rank} AB \leq \operatorname{rank} A$. (Remark: Here, one can use the fact that if $V \subset W$, then $\dim V \leq \dim W$. Do you understand why it is true?)

Answer. We have that $AV \subset AX$, and that $AX = \operatorname{Ran} A$. Thus, by the hint, since $AV \subset \operatorname{Ran} A$, we have that $\dim AV \leq \dim \operatorname{Ran} A$. But this implies that $\dim AV \leq \operatorname{rank} A$, as desired.

The column space of B will be a subspace of X. Additionally, we naturally have that $\operatorname{Ran} AB = A \cdot C(B)$, where $C(B)$ is the colunn space of B ($AB\mathbf{x} \in A \cdot C(B)$ since $B\mathbf{x} \in C(B)$ and vice versa). Thus, by the previous result, $\operatorname{rank} AB = \dim \operatorname{Ran} AB = \dim A \cdot C(B) \leq \operatorname{rank} A$, as desired. □

7.6. Prove that if the product AB of two $n \times n$ matrices is invertible, then both A and B are invertible. Even if you know about determinants, do not use them (we did not cover them yet). (Hint: Use the previous 2 problems.)

Answer. If AB is invertible, then it has a pivot in every column and row. Thus, $\operatorname{rank} AB = n$. It follows by Problem 7.4 that $n = \operatorname{rank} AB \leq \operatorname{rank} A \leq n$, implying that $\operatorname{rank} A = n$. Similarly, Problem 7.5 implies that $\operatorname{rank} B = n$. But these two results imply that A and B both have pivots in every colunn and row, i.e., both are invertible. □

7.9. If A has the same four fundamental subspaces as B, does $A = B$?

Answer. No — consider the following two matrices.

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad B = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

Both of these matrices have

$$\operatorname{Ker} X = \{\mathbf{0}\} \qquad \operatorname{Ran} X = \mathbb{R}^2 \qquad \operatorname{Ker} X^T = \{\mathbf{0}\} \qquad \operatorname{Ran} X^T = \mathbb{R}^2$$

where $X = A$ or B. However, we also clearly have $A \neq B$. □

7.14. Is it possible for a real matrix A that $\operatorname{Ran} A = \operatorname{Ker} A^T$? Is it possible for a complex A?

Answer. Suppose for the sake of contradiction that for a real $m \times n$ matrix $A : V \to W$, $\operatorname{Ran} A = \operatorname{Ker} A^T$. Then $A\mathbf{v} \in \operatorname{Ran} A = \operatorname{Ker} A^T$ for all $\mathbf{v} \in V$. It follows that $A^T(A\mathbf{v}) = \mathbf{0}$ for all $\mathbf{v} \in V$. Thus, $A^T A = 0$. Consequently,

$$\begin{aligned}0 &= \operatorname{tr}(0)\\ &= \operatorname{tr}(A^T A)\\ &= \sum_{j=1}^{n}(A^T A)_{ii}\\ &= \sum_{j=1}^{n}\sum_{i=1}^{m} A_{ij}^2\end{aligned}$$

It follows that $A_{ij} = 0$ for all i, j, i.e., that $A = 0$. But this implies that $\operatorname{Ran} A = \{\mathbf{0}\} \neq W = \operatorname{Ker} A^T$, a contradiction.

It is possible for a complex matrix: Consider

$$A = \begin{pmatrix} 0 & 1 \\ 0 & i \end{pmatrix}$$

Clearly

$$\operatorname{Ran} A = \operatorname{span}\begin{pmatrix} 1 \\ i \end{pmatrix}$$

and it can be shown that $\operatorname{Ker} A^T$ is the same. ☐

8.3. Find the change of coordinates matrix that changes the coordinates in the basis $1, 1+t$ in $\mathbb{P}_1$ to the coordinates in the basis $1-t, 2t$.

Answer. Let $\mathcal{A} = \{1, 1+t\}$, $\mathcal{B} = \{1-t, 2t\}$, and $\mathcal{S} = \{1, t\}$. Then following the procedure from Treil (2017), we have that

$$[I]_{\mathcal{S}\mathcal{A}} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \qquad [I]_{\mathcal{S}\mathcal{B}} = \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix}$$

so

$$\begin{aligned}[I]_{\mathcal{B}\mathcal{A}} &= [I]_{\mathcal{B}\mathcal{S}}[I]_{\mathcal{S}\mathcal{A}}\\ &= ([I]_{\mathcal{S}\mathcal{B}})^{-1}[I]_{\mathcal{S}\mathcal{A}}\\ &= \frac{1}{2}\begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\\ &= \frac{1}{2}\begin{pmatrix} 2 & 2 \\ 1 & 2 \end{pmatrix}\end{aligned}$$

☐

8.6. Are the matrices $\left(\begin{smallmatrix} 1 & 3 \\ 2 & 2 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 0 & 2 \\ 4 & 2 \end{smallmatrix}\right)$ similar? Justify.

Answer. We will first prove that if A and B are similar, then $\operatorname{tr}(A) = \operatorname{tr}(B)$. Let A, B be similar. Then $A = Q^{-1}BQ$, so

$$\begin{aligned}\operatorname{tr}(A) &= \operatorname{tr}(Q^{-1}BQ)\\ &= \operatorname{tr}(Q^{-1}QB)\\ &= \operatorname{tr}(B)\end{aligned}$$

as desired.

Now observe that $\operatorname{tr}\left(\begin{smallmatrix} 1 & 3 \\ 2 & 2 \end{smallmatrix}\right) = 3$ while $\operatorname{tr}\left(\begin{smallmatrix} 0 & 2 \\ 4 & 2 \end{smallmatrix}\right) = 2$. Thus, by the contrapositive of the lemma, we have that the two matrices aren't similar. ☐

2 Eigenvalues and Eigenvectors

From Treil (2017).

Chapter 4

10/11: **1.1.** True or false:

a) Every linear operator in an n-dimensional vector space has n distinct eigenvalues.

Answer. False.
The identity linear operator I_2 in $\mathbb{R}^2$ has the sole eigenvalue $\lambda = 1$, since $I_2\mathbf{x} = 1\mathbf{x}$ for any $\mathbf{x} \in \mathbb{R}^2$. ☐

b) If a matrix has one eigenvector, it has infinitely many eigenvectors.

Answer. True.
Let $A\mathbf{x} = \lambda\mathbf{x}$. Then $\alpha\mathbf{x}$ is also an eigenvector of A for any $\alpha \in \mathbb{F}$ since

$$A(\alpha\mathbf{x}) = \alpha A\mathbf{x} = \alpha\lambda\mathbf{x} = \lambda(\alpha\mathbf{x})$$

☐

c) There exists a square real matrix with no real eigenvalues.

Answer. True.
Consider

$$\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}$$

for which we have $\lambda = 1 \pm 2i$. Since the two eigenvalues $1 + 2i$ and $1 - 2i$ are distinct, and the square matrix given is 2×2, there are no more eigenvalues. Therefore, every eigenvalue of this matrix is not real. ☐

d) There exists a square matrix with no (complex) eigenvectors.

Answer. False.
Let $\mathbf{x}$ be an eigenvector of A. If $\mathbf{x}$ is complex, then we are done. If $\mathbf{x}$ is real, then multiply $\mathbf{x}$ by the scalar i. It follows by the proof of part (b) that $i\mathbf{x}$ is an eigenvector if A. ☐

e) Similar matrices always have the same eigenvalues.

Answer. True.
The characteristic polynomials of similar matrices coincide. ☐

f) Similar matrices always have the same eigenvectors.

Answer. False.
The matrix

$$\begin{pmatrix} 1 & 2 \\ 8 & 1 \end{pmatrix}$$

has eigenvectors

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix} \qquad \begin{pmatrix} 1 \\ -2 \end{pmatrix}$$

while its similar matrix

$$\begin{pmatrix} 5 & 0 \\ 0 & -3 \end{pmatrix}$$

has eigenvectors

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Note that since similar matrices refer to the same linear transformation, a single linear transformation technically only has one set of eigenvectors (albeit possibly expressed in different bases). □

g) A non-zero sum of two eigenvectors of a matrix A is always an eigenvector.

Answer. False.
Consider

$$\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

with eigenvalues $\lambda = 1, 2$ and respective eigenvectors

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Note that

$$\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

where the "**b**" vector is not a scalar multiple of the "**x**" vector. □

h) A non-zero sum of two eigenvectors of a matrix A corresponding to the same eigenvalue λ is always an eigenvector.

Answer. True.
Let $A\mathbf{x} = \lambda\mathbf{x}$ and $A\mathbf{y} = \lambda\mathbf{y}$. Then

$$\begin{aligned} A(\alpha\mathbf{x} + \beta\mathbf{y}) &= \alpha A\mathbf{x} + \beta A\mathbf{y} \\ &= \alpha\lambda\mathbf{x} + \beta\lambda\mathbf{y} \\ &= \lambda(\alpha\mathbf{x} + \beta\mathbf{y}) \end{aligned}$$

as desired. □

1.3. Compute eigenvalues and eigenvectors of the rotation matrix

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

Note that the eigenvalues (and eigenvectors) do not need to be real.

Answer. The characteristic polynomial of $A - \lambda I$ is

$$\begin{aligned} 0 &= \det(A - \lambda I) \\ &= (\cos\alpha - \lambda)^2 + \sin^2\alpha \\ -\sin^2\alpha &= (\cos\alpha - \lambda)^2 \\ \pm i\sin\alpha &= \pm\cos\alpha - \lambda \\ \lambda &= \cos\alpha + i\sin\alpha = \mathrm{e}^{i\alpha} \\ &= \cos\alpha - i\sin\alpha = \mathrm{e}^{-i\alpha} \end{aligned}$$

Thus, $\lambda = \mathrm{e}^{i\alpha}, \mathrm{e}^{-i\alpha}$. It follows by solving the systems of equations

$$\begin{aligned} x_1\cos\alpha - x_2\sin\alpha &= \mathrm{e}^{i\alpha}x_1 \\ x_1\sin\alpha + x_2\cos\alpha &= \mathrm{e}^{i\alpha}x_2 \end{aligned} \qquad \begin{aligned} y_1\cos\alpha - y_2\sin\alpha &= \mathrm{e}^{-i\alpha}y_1 \\ y_1\sin\alpha + y_2\cos\alpha &= \mathrm{e}^{-i\alpha}y_2 \end{aligned}$$

that the eigenvectors are

$$x = \begin{pmatrix} 1 \\ -i \end{pmatrix} \qquad\qquad y = \begin{pmatrix} 1 \\ i \end{pmatrix}$$

□

1.5. Prove that eigenvalues (counting multiplicities) of a triangular matrix coincide with its diagonal entries.

Answer. Since the determinant of a triangular matrix is the product of its diagonal entries, we have that

$$\det(A - \lambda I) = (a_{1,1} - \lambda)(a_{2,2} - \lambda) \cdots (a_{n,n} - \lambda)$$

But this polynomial is zero only if and only if λ is a diagonal entry, so the eigenvalues must be the diagonal entries. □

1.6. An operator A is called **nilpotent** if $A^k = \mathbf{0}$ for some k. Prove that if A is nilpotent, then $\sigma(A) = \{0\}$ (i.e., that 0 is the only eigenvalue of A).

Answer. Suppose for the sake of contradiction that λ is a nonzero eigenvalue of A with corresponding eigenvector $\mathbf{x}$. Then since $A\mathbf{x} = \lambda\mathbf{x}$, $A^k\mathbf{x} = \lambda^k\mathbf{x} \neq \mathbf{0} = 0\mathbf{x}$, so $A^k \neq 0$, a contradiction. □

1.7. Show that the characteristic polynomial of a block triangular matrix

$$\begin{pmatrix} A & * \\ \mathbf{0} & B \end{pmatrix}$$

where A and B are square matrices coincides with $\det(A - \lambda I)\det(B - \lambda I)$. (Hint: Use Exercise 3.11 from Chapter 3.)

Answer. It follows from Chapter 3, Exercise 3.11 that

$$\begin{aligned} \det\left(\begin{pmatrix} A & * \\ \mathbf{0} & B \end{pmatrix} - \lambda I\right) &= \det\begin{pmatrix} A - \lambda I & * \\ \mathbf{0} & B - \lambda I \end{pmatrix} \\ &= \det(A - \lambda I)\det(B - \lambda I) \end{aligned}$$

as desired. □

1.8. Let $\mathbf{v}_1, \dots, \mathbf{v}_n$ be a basis in a vector space V. Assume also that the first k vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$ of the basis are eigenvectors of an operator A, corresponding to an eigenvalue λ (i.e., that $A\mathbf{v}_j = \lambda\mathbf{v}_j$, $j = 1, \dots, k$). Show that in this basis, the matrix of the operator A has block triangular form

$$\begin{pmatrix} \lambda I_k & * \\ \mathbf{0} & B \end{pmatrix}$$

where I_k is the $k \times k$ identity matrix and B is some $(n-k) \times (n-k)$ matrix.

Answer. We will first show that if $\mathbf{v}_i$ is an eigenvector of A *and* a part of the basis $\mathbf{v}_1, \dots, \mathbf{v}_n$ of V, then it's matrix with respect to $\mathbf{v}_1, \dots, \mathbf{v}_n$ has zeros in every slot except the i^{th} slot, which is 1. This is easily shown as follows.

$$\begin{aligned} A\mathbf{v}_i &= \lambda\mathbf{v}_i \\ A\mathbf{v}_i &= \lambda(0\mathbf{v}_1 + \cdots + 0\mathbf{v}_{i-1} + 1\mathbf{v}_i + 0\mathbf{v}_{i+1} + \cdots + 0\mathbf{v}_n) \end{aligned}$$

This combined with the observations that the i^{th} column of A is equal to $A\mathbf{v}_i$ and $A\mathbf{v}_i = \lambda\mathbf{v}_i$ proves that

$$A = \begin{pmatrix} A\mathbf{v}_1 & \cdots & A\mathbf{v}_n \end{pmatrix} = \begin{pmatrix} \lambda\mathbf{v}_1 & \cdots & \lambda\mathbf{v}_k & A\mathbf{v}_{k+1} & \cdots & A\mathbf{v}_n \end{pmatrix} = \begin{pmatrix} \lambda I_k & * \\ \mathbf{0} & B \end{pmatrix}$$

as desired. □

1.10. Prove that the determinant of a matrix A is the product of its eigenvalues (counting multiplicities). (Hint: First show that $\det(A-\lambda I) = (\lambda_1 - \lambda)\cdots(\lambda_n - \lambda)$, where $\lambda_1, \dots, \lambda_n$ are eigenvalues (counting multiplicities). Then compare the free terms (terms without λ) or plug in $\lambda = 0$ to get the conclusion.)

Answer. We know that the roots of the characteristic polynomial $\det(A-\lambda I)$ of A are exactly the eigenvalues $\lambda_1, \dots, \lambda_n$ of A. In other words, $\det(A-\lambda I)$ must go to zero exactly when $\lambda = \lambda_i$ for some i. Thus, $\det(A-\lambda I)$ must be of the form

$$c(\lambda_1 - \lambda)\cdots(\lambda_n - \lambda)$$

for some $c \in \mathbb{F}$. But since λ only occurs in $A - \lambda I$ with the coefficient -1, and the λ^n term is solely generated by the term in the permutation sum that is the product of the diagonal entries, the λ^n term must have coefficient $(-1)^n$. Additionally, the polynomial above will have λ^n have coefficient $(-1)^n$. Thus, we must have $c = 1$, and we have proven the hint. Therefore,

$$\begin{aligned}\det A &= \det(A - 0I)\\ &= (\lambda_1 - 0)\cdots(\lambda_n - 0)\\ &= \lambda_1 \cdots \lambda_n\end{aligned}$$

as desired. □

1.11. Prove that the trace of a matrix equals the sum of its eigenvalues in three steps. First, compute the coefficient of λ^{n-1} in the right side of the equality

$$\det(A-\lambda I) = (\lambda_1 - \lambda)\cdots(\lambda_n - \lambda)$$

Then show that $\det(A-\lambda I)$ can be represented as

$$\det(A-\lambda I) = (a_{1,1} - \lambda)(a_{2,2} - \lambda)\cdots(a_{n,n} - \lambda) + q(\lambda)$$

where $q(\lambda)$ is a polynomial of degree at most $n-2$. And finally, compare the coefficients of λ^{n-1} to get the conclusion.

Answer. Consider $(\lambda_1 - \lambda)\cdots(\lambda_n - \lambda)$. We have that every λ^{n-1} term in the expansion of this product must take the λ from $n-1$ of the terms and the λ_i from the remaining term. Thus, our expansion should contain the terms $\lambda_1\lambda^{n-1}, \dots, \lambda_n\lambda^{n-1}$, which, when we sum, gives $(-1)^n(\lambda_1 + \cdots + \lambda_n)\lambda^{n-1}$.

In the permutation sum form of the determinant, we have that $(a_{1,1} - \lambda)(a_{2,2} - \lambda)\cdots(a_{n,n} - \lambda)$ will be one of the terms in the sum. In particular, it is the *only* term to contain all the λ-containing entries in the matrix, so it solely determines the λ^n term. Additionally, the term containing the next-highest number of λ's must contain $n-2$ λ's, not $n-1$, since any product with $n-1$ diagonal entries and 1 non-diagonal entry necessarily contains two terms that are in the same row or column. Thus, the term given solely determines the λ^{n-1} term as well. All of the other terms, having degree at most λ^{n-2}, can be defined equal to $q(\lambda)$.

Therefore, since the first part of the proof gives

$$(\lambda_1 + \cdots + \lambda_n)\lambda^{n-1}$$

as the λ^{n-1} term, and the second part of the proof (by a similar argument) gives

$$(a_{1,1} + a_{2,2} + \cdots + a_{n,n})\lambda^{n-1}$$

as the λ^{n-1} term, we have by comparing terms (rigorously, subtract all terms of other degrees to preserve the equality) that

$$\operatorname{tr} A = a_{1,1} + a_{2,2} + \cdots + a_{n,n} = \lambda_1 + \cdots + \lambda_n$$

as desired. □

2.1. Let A be an $n \times n$ matrix. True or false (justify your conclusions):

a) A^T has the same eigenvalues as A.

Answer. True.
Since $\det B = \det B^T$ for any matrix B and the transpose operation does not affect the diagonal, we have that

$$\begin{aligned}\det(A - \lambda I) &= \det\big((A - \lambda I)^T\big) \\ &= \det\big(A^T - (\lambda I)^T\big) \\ &= \det\big(A^T - \lambda I\big)\end{aligned}$$

as desired. □

b) A^T has the same eigenvectors as A.

Answer. False.
Let

$$A = \begin{pmatrix} 1 & 2 \\ 8 & 1 \end{pmatrix}$$

Then we can calculate that A has eigenvectors

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix} \qquad \begin{pmatrix} -1 \\ 2 \end{pmatrix}$$

but A^T has eigenvectors

$$\begin{pmatrix} 2 \\ 1 \end{pmatrix} \qquad \begin{pmatrix} -2 \\ 1 \end{pmatrix}$$

□

c) If A is diagonalizable, then so is A^T.

Answer. True.
Suppose $A = SDS^{-1}$. Then

$$\begin{aligned}A^T &= (SDS^{-1})^T \\ &= (S^{-1})^T D^T S^T \\ &= (S^{-1})^T D((S^{-1})^T)^{-1}\end{aligned}$$

as desired. □

2.2. Let A be a square matrix with real entries, and let λ be its complex eigenvalue. Suppose $\mathbf{v} = (v_1, \dots, v_n)^T$ is a corresponding eigenvector, i.e., $A\mathbf{v} = \lambda\mathbf{v}$. Prove that the $\bar{\lambda}$ is an eigenvalue of A and $A\bar{\mathbf{v}} = \bar{\lambda}\bar{\mathbf{v}}$, where $\bar{\mathbf{v}} = (\bar{v}_1, \dots, \bar{v}_n)^T$ is the complex conjugate of the vector $\mathbf{v}$.

Answer. Let $\mathbf{v} = \mathbf{a} + i\mathbf{b}$ where $a_j = \operatorname{Re} v_j$ and $b_j = \operatorname{Im} v_j$. It follows that

$$A\mathbf{a} + iA\mathbf{b} = A\mathbf{v} = \lambda\mathbf{v} = \lambda\mathbf{a} + i\lambda\mathbf{b}$$

This combined with the fact that all entries in $A, \mathbf{a}, \mathbf{b}$ are real implies by matching corresponding parts that

$$A\mathbf{a} = \lambda\mathbf{a} \qquad A\mathbf{b} = \lambda\mathbf{b}$$

Therefore,

$$A\bar{\mathbf{v}} = A(\mathbf{a} - i\mathbf{b}) = A\mathbf{a} - iA\mathbf{b} = \lambda\mathbf{a} - i\lambda\mathbf{b} = \lambda(\mathbf{a} - i\mathbf{b}) = \lambda\bar{\mathbf{v}}$$

as desired. □

2.3. Let

$$A = \begin{pmatrix} 4 & 3 \\ 1 & 2 \end{pmatrix}$$

Find A^{2004} by diagonalizing A.

Answer. We have that

$$\begin{aligned} 0 &= \det(A - \lambda I) \\ &= (4-\lambda)(2-\lambda) - 3 \\ &= \lambda^2 - 6\lambda + 5 \\ &= (\lambda - 5)(\lambda - 1) \end{aligned}$$

Thus, $\lambda = 5, 1$. It follows by inspection that

$$x_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \qquad x_2 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$$

Consequently,

$$S = \begin{pmatrix} -1 & 3 \\ 1 & 1 \end{pmatrix} \qquad S^{-1} = \frac{1}{4}\begin{pmatrix} -1 & 3 \\ 1 & 1 \end{pmatrix}$$

Hence

$$A = \frac{1}{4}\begin{pmatrix} -1 & 3 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 5 \end{pmatrix}\begin{pmatrix} -1 & 3 \\ 1 & 1 \end{pmatrix}$$

Therefore,

$$\begin{aligned} A^{2004} &= \frac{1}{4}\begin{pmatrix} -1 & 3 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 5^{2004} \end{pmatrix}\begin{pmatrix} -1 & 3 \\ 1 & 1 \end{pmatrix} \\ &= \frac{1}{4}\begin{pmatrix} -1 & 3 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} -1 & 3 \\ 5^{2004} & 5^{2004} \end{pmatrix} \\ &= \frac{1}{4}\begin{pmatrix} 1 + 3 \cdot 5^{2004} & -3 + 3 \cdot 5^{2004} \\ -1 + 5^{2004} & 3 + 5^{2004} \end{pmatrix} \end{aligned}$$

□

2.4. Construct a matrix A with eigenvalues 1 and 3 and corresponding eigenvectors $(1,2)^T$ and $(1,1)^T$. Is such a matrix unique?

Answer. Let

$$\begin{aligned} A &= \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} -1 & 1 \\ 2 & -1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} -1 & 1 \\ 6 & -3 \end{pmatrix} \\ &= \begin{pmatrix} 5 & -2 \\ 4 & -1 \end{pmatrix} \end{aligned}$$

Suppose A' has eigenvalues $1, 3$ with corresponding eigenvectors $(1,2)^T$ and $(1,1)^T$. Then since the eigenvectors are linearly independent and form a basis of $\mathbb{R}^2$, Theorem 2.1 implies that A' is diagonal with diagonal matrix equal to the middle matrix in the first line above and change of basis matrices equal to the other two in the first line above. Therefore, $A = A'$. □

2.6. Consider the matrix
$$A = \begin{pmatrix} 2 & 6 & -6 \\ 0 & 5 & -2 \\ 0 & 0 & 4 \end{pmatrix}$$

a) Find its eigenvalues. Is it possible to find the eigenvalues without computing?

Answer. It's eigenvalues are $\lambda = 2, 5, 4$, since this is an upper-triangular matrix and those are the diagonal entries. □

b) Is this matrix diagonalizable? Find out without computing anything.

Answer. Yes. Since the eigenvalues are all distinct and there are 3 for this 3×3 matrix, Corollary 2.3 implies that A is diagonalizable. □

c) If the matrix is diagonalizable, diagonalize it.

Answer. If $\lambda_1 = 2$, $\lambda_2 = 5$, and $\lambda_3 = 4$, then the corresponding eigenvectors are
$$x_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad x_2 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} \qquad x_3 = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$$

It follows that
$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} 1 & -2 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$$

□

2.8. Find all square roots of the matrix
$$A = \begin{pmatrix} 5 & 2 \\ -3 & 0 \end{pmatrix}$$
i.e., find all matrices B such that $B^2 = A$. (Hint: Finding a square root of a diagonal matrix is easy. You can leave your answer as a product.)

Answer. We have that
$$A = \begin{pmatrix} -1 & -2 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} -3 & -2 \\ 1 & 1 \end{pmatrix}$$
Therefore, we have four possibilities for B:
$$B_1 = \begin{pmatrix} -1 & -2 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} \sqrt{3} & 0 \\ 0 & \sqrt{2} \end{pmatrix} \begin{pmatrix} -3 & -2 \\ 1 & 1 \end{pmatrix}$$
$$B_2 = \begin{pmatrix} -1 & -2 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} -\sqrt{3} & 0 \\ 0 & \sqrt{2} \end{pmatrix} \begin{pmatrix} -3 & -2 \\ 1 & 1 \end{pmatrix}$$
$$B_3 = \begin{pmatrix} -1 & -2 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} \sqrt{3} & 0 \\ 0 & -\sqrt{2} \end{pmatrix} \begin{pmatrix} -3 & -2 \\ 1 & 1 \end{pmatrix}$$
$$B_4 = \begin{pmatrix} -1 & -2 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} -\sqrt{3} & 0 \\ 0 & -\sqrt{2} \end{pmatrix} \begin{pmatrix} -3 & -2 \\ 1 & 1 \end{pmatrix}$$

□

2.10. Let A be a 5×5 matrix with 3 eigenvalues (not counting multiplicities). Suppose we know that one eigenspace is three-dimensional. Can you say if A is diagonalizable?

Answer. Yes, it is diagonalizable. Let $\lambda_1, \lambda_2, \lambda_3$ be the 3 eigenvalues of A, let $\mathbf{v}_1, \mathbf{v}_2$ be the eigenvectors corresponding to λ_1, λ_2, and let $\mathbf{v}_{3a}, \mathbf{v}_{3b}, \mathbf{v}_{3c}$ be a basis of the eigenvectors corresponding to λ_3. Since the eigenspace of λ_3 is three dimensional, we know that $\mathbf{v}_{3a}, \mathbf{v}_{3b}, \mathbf{v}_{3c}$ is linearly independent. Additionally, we have by consecutive applications of Theorem 2.2 that $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_{3a}$, $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_{3b}$, and $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_{3c}$ are linearly independent lists. Hence $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_{3a}, \mathbf{v}_{3b}, \mathbf{v}_{3c}$ is a linearly independent list of length 5, so it must form a basis of $\mathbb{F}^5$. Therefore, by Theorem 2.1, A is diagonalizable. □

2.11. Give an example of a 3×3 matrix which cannot be diagonalized. After you construct the matrix, can you make it "generic," so no special structure of the matrix can be seen?

Answer. Generalizing from the given example, we can show that

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

is not diagonalizable. Applying row operations can put the matrix in the more generic form

$$\begin{pmatrix} 0 & 1 & 4 \\ 1 & 2 & 3 \\ -1 & 0 & 4 \end{pmatrix}$$

□

2.13. Eigenvalues of a transposition:

a) Consider the transformation T in the space $M_{2\times 2}$ of 2×2 matrices defined by $T(A) = A^T$. Find all its eigenvalues and eigenvectors. Is it possible to diagonalize this transformation? (Hint: While it is possible to write a matrix of this linear transformation in some basis, compute the characteristic polynomial, and so on, it is easier to find eigenvalues and eigenvectors directly from the definition.)

Answer. The symmetric matrices are eigenvectors of this transformation with eigenvalue 1. A basis of them would be

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

The antisymmetric matrices are eigenvectors of this transformation with eigenvalue -1. A basis of them would be

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

Since these four matrices are linearly independent, there exists a basis of $M_{2\times 2}$ of eigenvectors of T. Therefore, T is diagonalizable. □

b) Can you do the same problem but in the space of $n \times n$ matrices?

Answer. Yes. A basis of the $n \times n$ symmetric matrices includes all of the matrices that are zero everywhere except for one 1 in a diagonal entry, and all of the matrices that are zero everywhere except for two 1's in off-diagonal symmetric positions. There are $\frac{n}{2}(n+1)$ of these basis "vectors." A basis of the $n \times n$ antisymmetric matrices includes all of the matrices that are zero everywhere except for a -1 in an off-diagonal position in the upper triangle and a 1 in the symmetric position in the lower triangle. There are $\frac{n}{2}(n-1)$ of these. Together, we have

$$\frac{n}{2}(n+1) + \frac{n}{2}(n-1) = n^2$$

basis "vectors," meaning that we have a complete eigenbasis of $M_{n\times n}$. □

2.14. Prove that two subspaces V_1 and V_2 are linearly independent if and only if $V_1 \cap V_2 = \{\mathbf{0}\}$.

Answer. Suppose first that V_1, V_2 are linearly independent. Let $\mathbf{v}_{11}, \ldots, \mathbf{v}_{1n}$ be a basis of V_1, and let $\mathbf{v}_{21}, \ldots, \mathbf{v}_{2m}$ be a basis of V_2. Then by Lemma 2.7, $\mathbf{v}_{11}, \ldots, \mathbf{v}_{1n}, \mathbf{v}_{21}, \ldots, \mathbf{v}_{2m}$ is linearly independent. Now suppose $\mathbf{v} \in V_1 \cap V_2$. Since $\mathbf{v} \in V_1$, $\mathbf{v} = \alpha_{11}\mathbf{v}_{11} + \cdots + \alpha_{1n}\mathbf{v}_{1n}$. Similarly, $\mathbf{v} = \alpha_{21}\mathbf{v}_{21} + \cdots + \alpha_{2m}\mathbf{v}_{2m}$. Thus,

$$\mathbf{0} = \mathbf{v} - \mathbf{v} = \alpha_{11}\mathbf{v}_{11} + \cdots + \alpha_{1n}\mathbf{v}_{1n} - \alpha_{21}\mathbf{v}_{21} - \cdots - \alpha_{2m}\mathbf{v}_{2m}$$

But since $\mathbf{v}_{11}, \ldots, \mathbf{v}_{1n}, \mathbf{v}_{21}, \ldots, \mathbf{v}_{2m}$ is linearly independent, it follows that all the α's are 0. Therefore, $\mathbf{v} = 0\mathbf{v}_{11} + \cdots + 0\mathbf{v}_{1n} = \mathbf{0}$, so $V_1 \cap V_2 \subset \{\mathbf{0}\}$. The inclusion in the other direction is obvious, since V_1, V_2 are subspaces.

Now suppose that $V_1 \cap V_2 = \{\mathbf{0}\}$. To prove that V_1, V_2 are linearly independent, it will suffice to show that $\mathbf{v}_1 + \mathbf{v}_2 = \mathbf{0}$ where $\mathbf{v}_i \in V_i$ for all i implies $\mathbf{v}_i = \mathbf{0}$ for all i. Let $\mathbf{v}_1 + \mathbf{v}_2 = \mathbf{0}$ where $\mathbf{v}_i \in V_i$ for all i. Suppose for the sake of contradiction that $\mathbf{v}_1 \neq \mathbf{0}$. Then we must have $\mathbf{v}_2 = -\mathbf{v}_1 \neq \mathbf{0}$. But by closure under scalar multiplication, this implies that $-1 \cdot -\mathbf{v}_1 = \mathbf{v}_1 \in V_2$ since $\mathbf{v}_2 \in V_2$. Therefore, $\mathbf{v}_1 \in V_1 \cap V_2$ as well, a contradiction. The proof is symmetric if we let $\mathbf{v}_2 \neq \mathbf{0}$ first. $\square$

3 Inner Product Spaces

From Treil (2017).

Chapter 5

10/18: **3.2.** Apply Gram-Schmidt orthogonalization to the system of vectors $(1,2,3)^T$, $(1,3,1)^T$. Write the matrix of the orthogonal projection onto the 2-dimensional subspace spanned by these vectors.

Answer. We define

$$\mathbf{v}_1 = \begin{pmatrix}1\\2\\3\end{pmatrix} \qquad \begin{aligned}\mathbf{v}_2 &= \begin{pmatrix}1\\3\\1\end{pmatrix} - \frac{((1,3,1)^T,(1,2,3)^T)}{\|(1,2,3)^T\|^2}\begin{pmatrix}1\\2\\3\end{pmatrix}\\ &= \begin{pmatrix}1\\3\\1\end{pmatrix} - \frac{10}{14}\begin{pmatrix}1\\2\\3\end{pmatrix}\\ &= \frac{1}{7}\begin{pmatrix}2\\11\\-8\end{pmatrix}\end{aligned}$$

Thus, we have that

$$\begin{aligned}P_{\{\mathbf{v}_1,\mathbf{v}_2\}} &= \sum_{k=1}^{2}\frac{1}{\|\mathbf{v}_k\|^2}\mathbf{v}_k\mathbf{v}_k^*\\ &= \frac{1}{1^2+2^2+3^2}\begin{pmatrix}1\\2\\3\end{pmatrix}\begin{pmatrix}1 & 2 & 3\end{pmatrix} + \frac{1}{2^2+11^2+(-8)^2}\begin{pmatrix}2\\11\\-8\end{pmatrix}\begin{pmatrix}2 & 11 & -8\end{pmatrix}\\ &= \frac{1}{14}\begin{pmatrix}1 & 2 & 3\\2 & 4 & 6\\3 & 6 & 9\end{pmatrix} + \frac{1}{189}\begin{pmatrix}4 & 22 & -16\\22 & 121 & -88\\-16 & -88 & 64\end{pmatrix}\\ &= \frac{1}{54}\begin{pmatrix}5 & 14 & 7\\14 & 50 & -2\\7 & -2 & 53\end{pmatrix}\end{aligned}$$

□

3.5. Find the orthogonal projection of a vector $(1,1,1,1)^T$ onto the subspace spanned by the vectors $\mathbf{v}_1 = (1,3,1,1)^T$ and $\mathbf{v}_2 = (2,-1,1,0)^T$ (note that $\mathbf{v}_1 \perp \mathbf{v}_2$).

Answer. If $\mathbf{v} = (1,1,1,1)^T$,

$$\begin{aligned}P_{\{\mathbf{v}_1,\mathbf{v}_2\}}\mathbf{v} &= \sum_{k=1}^{2}\frac{(\mathbf{v},\mathbf{v}_k)}{\|\mathbf{v}_k\|^2}\mathbf{v}_k\\ &= \frac{((1,1,1,1)^T,(1,3,1,1)^T)}{1^2+3^2+1^2+1^2}\begin{pmatrix}1\\3\\1\\1\end{pmatrix} + \frac{((1,1,1,1)^T,(2,-1,1,0)^T)}{2^2+(-1)^2+1^2+0^2}\begin{pmatrix}2\\-1\\1\\0\end{pmatrix}\\ &= \frac{6}{12}\begin{pmatrix}1\\3\\1\\1\end{pmatrix} + \frac{2}{6}\begin{pmatrix}2\\-1\\1\\0\end{pmatrix}\end{aligned}$$

$$= \frac{1}{6}\begin{pmatrix} 7 \\ 7 \\ 5 \\ 3 \end{pmatrix}$$

□

3.6. Find the distance from a vector $(1,2,3,4)^T$ to the subspace spanned by the vectors $\mathbf{v}_1 = (1,-1,1,0)^T$ and $\mathbf{v}_2 = (1,2,1,1)^T$ (note that $\mathbf{v}_1 \perp \mathbf{v}_2$). Can you find the distance without actually computing the projection? That would simplify the calculations.

Answer. Let $\mathbf{v} = (1,2,3,4)^T$. Suppose $\mathbf{v} = \mathbf{u} + \mathbf{w}$, where $\mathbf{u} \in \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2)$ and $\mathbf{w} \perp \operatorname{span}(\mathbf{v}_1, \mathbf{v}_2)$. Clearly $\mathbf{u} \perp \mathbf{w}$. Consequently, by the Pythagorean theorem,

$$\|\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{w}\|^2$$

where $\|\mathbf{w}\|$ is the desired distance from $\mathbf{v}$ to $\operatorname{span}(\mathbf{v}_1, \mathbf{v}_2)$. $\|\mathbf{v}\|$ is easy to find, but $\|\mathbf{u}\|$ presents a bit more of a challenge. However, since $\mathbf{v}_1 \perp \mathbf{v}_2$, we can project $\mathbf{v}$ onto $\mathbf{v}_1$ and $\mathbf{v}_2$ separately (an easier process than computing the whole projection), and know that

$$\|\mathbf{u}\|^2 = \|\alpha_1 \mathbf{v}_1\|^2 + \|\alpha_2 \mathbf{v}_2\|^2$$

Combining the above two equations, we have that

$$\|\mathbf{v}\|^2 = \|\alpha_1 \mathbf{v}_1\|^2 + \|\alpha_2 \mathbf{v}_2\|^2 + \|\mathbf{w}\|^2$$

But since $\alpha_k = (\mathbf{v}, \mathbf{v}_k)/\|\mathbf{v}_k\|^2$ for $k = 1, 2$, we have that

$$\begin{aligned}
\|\mathbf{w}\| &= \sqrt{\|\mathbf{v}\|^2 - \left\| \frac{(\mathbf{v}, \mathbf{v}_1)}{\|\mathbf{v}_1\|^2} \mathbf{v}_1 \right\|^2 - \left\| \frac{(\mathbf{v}, \mathbf{v}_2)}{\|\mathbf{v}_2\|^2} \mathbf{v}_2 \right\|^2} \\
&= \sqrt{\|\mathbf{v}\|^2 - \frac{(\mathbf{v}, \mathbf{v}_1)^2}{\|\mathbf{v}_1\|^2} - \frac{(\mathbf{v}, \mathbf{v}_2)^2}{\|\mathbf{v}_2\|^2}} \\
&= \sqrt{30 - \frac{2^2}{3} - \frac{12^2}{7}} \\
&= \sqrt{170/21}
\end{aligned}$$

□

3.7. True or false: If E is a subspace of V, then $\dim E + \dim(E^\perp) = \dim V$. Justify.

Answer. True.

Let E be a subspace of V with orthonormal basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$, and let $E^\perp$ have orthonormal basis $\mathbf{w}_1, \ldots, \mathbf{w}_m$. To prove that $\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{w}_1, \ldots, \mathbf{w}_n$ is a basis of V, it will suffice to show that the list is linearly independent and spanning. Let $k \in [m]$. Since $\mathbf{w}_k \in E^\perp$ and $\mathbf{v}_1, \ldots, vm_n \in E$, we have that $\mathbf{w}_k \perp \mathbf{v}_l$ for each $l = 1, \ldots, n$. Thus, by Corollary 2.6, $\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{w}_k$ is linearly independent. It follows by combining these k results that $\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{w}_1, \ldots, \mathbf{w}_m$ is linearly independent, as desired. On the other hand, by definition any vector $\mathbf{u} \in V$ admits a unique representation $\mathbf{u} = \mathbf{u}_1 + \mathbf{u}_2$ where $\mathbf{u}_1 \in E$ and $\mathbf{u}_2 \in E^\perp$. Additionally, $\mathbf{u}_1 = \sum_{k=1}^n \alpha_k \mathbf{v}_k$ for some $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$ and $\mathbf{u}_2 = \sum_{k=1}^m \beta \mathbf{w}_k$ for some $\beta_1, \ldots, \beta_n \in \mathbb{F}$. Thus, $\mathbf{u} = \mathbf{u}_1 + \mathbf{u}_2$ can be expressed as a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{w}_1, \ldots, \mathbf{w}_n$, so the list is spanning, as desired. □

3.8. Let P be the orthogonal projection onto a subspace E of an inner product space V, let $\dim V = n$, and let $\dim E = r$. Find the eigenvalues and the eigenvectors (eigenspaces). Find the algebraic and geometric multiplicities of each eigenvalue.

Answer. The eigenvalues are 1 and 0, with corresponding eigenspaces E and $E^{\perp}$. It follows that the geometric multiplicity of 1 is r, and the geometric multiplicity of 0 is $n-r$. We will now prove that the algebraic multiplicities are the same. First off, Proposition 1.1 (Chapter 4) states that the algebraic multiplicity is greater than or equal to the geometric multiplicity. It follows that the algebraic multiplicity of 1 is greater than or equal to r and the algebraic multiplicity of 0 is greater than or equal to $n-r$. But since the total multiplicity cannot exceed n and $n-r+r=n$, we must have that the algebraic multiplicity of 1 is *equal* to r, and likewise for 0 and $n-r$. $\square$

3.9. Using eigenvalues to compute determinants:

a) Find the matrix of the orthogonal projection onto the one-dimensional subspace in $\mathbb{R}^n$ spanned by the vector $(1,\dots,1)^T$.

Answer. We have that

$$\begin{aligned} P_{\mathbf{v}} &= \frac{1}{\|\mathbf{v}\|^2}\mathbf{v}\mathbf{v}^* \\ &= \frac{1}{\sum_{i=1}^n 1^2}\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}\begin{pmatrix}1 & \cdots & 1\end{pmatrix} \\ &= \frac{1}{n}\begin{pmatrix}1 & \cdots & 1\\ \vdots & \ddots & \vdots\\ 1 & \cdots & 1\end{pmatrix} \end{aligned}$$

In other words, the desired matrix is the $n\times n$ matrix with $a_{i,j}=\frac{1}{n}$ for each $i,j=1,\dots,n$. $\square$

b) Let A be the $n\times n$ matrix with all entries equal to 1. Compute its eigenvalues and their multiplicities (use the previous problem).

Answer. We have that $A=nP_{\mathbf{v}}$ where $P_{\mathbf{v}}$ is the matrix defined in part (a). Since $P_{\mathbf{v}}$ is a projection matrix, the only two kinds of vectors that it *scales* are vectors in the subspace onto which it is projecting (these get mapped to themselves, i.e., scaled by 1) and vectors perpendicular to the subspace onto which it is projecting (these get scaled by 0 to $\mathbf{0}$). Thus, since $\dim\operatorname{span}\{\mathbf{v}\}=1$ and $\dim\operatorname{span}\{\mathbf{v}\}^{\perp}=n-1$, we have that

$$A\mathbf{x}=nP_{\mathbf{v}}\mathbf{x}=n\cdot 1\mathbf{x}=n\mathbf{x}$$

for all $\mathbf{x}\in\operatorname{span}\{\mathbf{v}\}$ and that

$$A\mathbf{x}=nP_{\mathbf{v}}\mathbf{x}=n\cdot 0\mathbf{x}=0\mathbf{x}$$

for all $\mathbf{x}\in\operatorname{span}\{\mathbf{v}\}^{\perp}$. Therefore, the eigenvalues of A are n and 0, with respective geometric and algebraic multiplicities 1 and $n-1$. $\square$

c) Compute eigenvalues (and multiplicities) of the matrix $A-I$, i.e., of the matrix with zeroes on the main diagonal and ones everywhere else.

Answer. Adapting from the part (c), if $A\mathbf{x}=n\mathbf{x}$, then

$$(A-I)\mathbf{x}=A\mathbf{x}-I\mathbf{x}=n\mathbf{x}-1\mathbf{x}=(n-1)\mathbf{x}$$

Similarly, if $A\mathbf{x}=0\mathbf{x}$, then

$$(A-I)\mathbf{x}=A\mathbf{x}-I\mathbf{x}=0\mathbf{x}-1\mathbf{x}=-1\mathbf{x}$$

Now suppose for the sake of contradiction that $-1\neq\lambda\neq n-1$ is an eigenvalue of $A-I$ with corresponding eigenvector $\mathbf{x}$. Then

$$A\mathbf{x}=(A-I)\mathbf{x}+I\mathbf{x}=\lambda\mathbf{x}+1\mathbf{x}=(\lambda+1)\mathbf{x}$$

so $\lambda+1$ is an eigenvalue of A. But this contradicts part (b). Therefore, $n-1$ and -1, with respective geometric and algebraic multiplicities 1 and $n-1$, are the only eigenvalues of $A-I$. $\square$

d) Compute $\det(A - I)$.

Answer. By part (b), $\det(A - \lambda I) = (n-\lambda)^1(0-\lambda)^{n-1}$. Thus,

$$\det(A - I) = (n-1)(-1)^{n-1}$$

as desired. □

3.11. Let $P = P_E$ be the matrix of an orthogonal projection onto a subspace E. Show that

a) The matrix P is self-adjoint, meaning that $P^* = P$.

Answer. Suppose p_{ij} is the entry in the i^{th} row and j^{th} column of P and $\mathbf{v}_1, \dots, \mathbf{v}_r$ is a basis of E. Let $\mathbf{v}_{k_i}$ denote the i^{th} coordinate of $\mathbf{v}_k$. Then we have that

$$\begin{aligned}
p_{ij} &= \sum_{k=1}^{r} \frac{\mathbf{v}_{k_i}\bar{\mathbf{v}}_{k_j}}{\|\mathbf{v}_k\|} \\
&= \sum_{k=1}^{r} \frac{\bar{\mathbf{v}}_{k_j}\bar{\bar{\mathbf{v}}}_{k_i}}{\|\mathbf{v}_k\|} \\
&= \sum_{k=1}^{r} \overline{\frac{\mathbf{v}_{k_j}\bar{\mathbf{v}}_{k_i}}{\|\mathbf{v}_k\|}} \\
&= \overline{\sum_{k=1}^{r} \frac{\mathbf{v}_{k_j}\bar{\mathbf{v}}_{k_i}}{\|\mathbf{v}_k\|}} \\
&= \bar{p}_{ji}
\end{aligned}$$

as desired. □

b) $P^2 = P$.

Answer. Let $\mathbf{v} = \mathbf{u} + \mathbf{w}$ where $\mathbf{u} \in E$ and $\mathbf{w} \in E^\perp$. Then

$$P^2\mathbf{v} = P(P\mathbf{v}) = P\mathbf{u} = \mathbf{u} = P\mathbf{v}$$

as desired. □

3.13. Suppose P is the orthogonal projection onto a subspace E, and Q is the orthogonal projection onto the orthogonal complement $E^\perp$.

a) What are $P + Q$ and PQ?

Answer. Let $\mathbf{v} = \mathbf{u} + \mathbf{w}$, where $\mathbf{u} \in E$ and $\mathbf{w} \in E^\perp$. Then

$$\begin{aligned}
(P+Q)\mathbf{v} &= P\mathbf{v} + Q\mathbf{v} & \qquad (PQ)\mathbf{v} &= P(Q\mathbf{v}) \\
&= \mathbf{u} + \mathbf{w} & &= P\mathbf{w} \\
&= \mathbf{v} & &= \mathbf{0} \\
&= I\mathbf{v} & &= 0\mathbf{v}
\end{aligned}$$

so $P + Q = I$ and $PQ = 0$. □

b) Show that $P - Q$ is its own inverse.

Answer. To show that $P - Q$ is its own inverse, it will suffice to show that $(P-Q)^2 = I$. Let $\mathbf{v} = \mathbf{u} + \mathbf{w}$ as in part (a). Then

$$\begin{aligned}
(P-Q)^2\mathbf{v} &= (P-Q)(P\mathbf{v} - Q\mathbf{v}) \\
&= (P-Q)(\mathbf{u} - \mathbf{w}) \\
&= P(\mathbf{u} - \mathbf{w}) - Q(\mathbf{u} - \mathbf{w}) \\
&= \mathbf{u} - (-\mathbf{w}) \\
&= \mathbf{v} \\
&= I\mathbf{v}
\end{aligned}$$

as desired. □

4.2. Find the matrix of the orthogonal projection P onto the column space of

$$\begin{pmatrix} 1 & 1 \\ 2 & -1 \\ -2 & 4 \end{pmatrix}$$

Use two methods: Gram-Schmidt orthogonalization and the formula for the projection. Compare the results.

Answer. Gram-Schmidt orthogonalization: We have that

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix} \qquad \begin{aligned}
\mathbf{v}_2 &= \begin{pmatrix} 1 \\ -1 \\ 4 \end{pmatrix} - \frac{((1,-1,4)^T, (1,2,-2)^T)}{\|(1,2,-2)^T\|^2} \begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix} \\
&= \begin{pmatrix} 1 \\ -1 \\ 4 \end{pmatrix} - \frac{-9}{9} \begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix} \\
&= \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}
\end{aligned}$$

so that

$$\begin{aligned}
P &= \sum_{k=1}^{2} \frac{1}{\|\mathbf{v}_k\|^2} \mathbf{v}_k \mathbf{v}_k^* \\
&= \frac{1}{1^2 + 2^2 + (-2)^2} \begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix} \begin{pmatrix} 1 & 2 & -2 \end{pmatrix} + \frac{1}{2^2 + 1^2 + 2^2} \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} \begin{pmatrix} 2 & 1 & 2 \end{pmatrix} \\
&= \frac{1}{9} \begin{pmatrix} 1 & 2 & -2 \\ 2 & 4 & -4 \\ -2 & -4 & 4 \end{pmatrix} + \frac{1}{9} \begin{pmatrix} 4 & 2 & 4 \\ 2 & 1 & 2 \\ 4 & 2 & 4 \end{pmatrix} \\
&= \frac{1}{9} \begin{pmatrix} 5 & 4 & 2 \\ 4 & 5 & -2 \\ 2 & -2 & 8 \end{pmatrix}
\end{aligned}$$

Formula: We have that

$$\begin{aligned}
P &= A(A^*A)^{-1}A^* \\
&= \begin{pmatrix} 1 & 1 \\ 2 & -1 \\ -2 & 4 \end{pmatrix} \left(\begin{pmatrix} 1 & 2 & -2 \\ 1 & -1 & 4 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & -1 \\ -2 & 4 \end{pmatrix} \right)^{-1} \begin{pmatrix} 1 & 2 & -2 \\ 1 & -1 & 4 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 1 \\ 2 & -1 \\ -2 & 4 \end{pmatrix} \begin{pmatrix} 9 & -9 \\ -9 & 18 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 2 & -2 \\ 1 & -1 & 4 \end{pmatrix} \\
&= \frac{1}{9} \begin{pmatrix} 1 & 1 \\ 2 & -1 \\ -2 & 4 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & -2 \\ 1 & -1 & 4 \end{pmatrix} \\
&= \frac{1}{9} \begin{pmatrix} 1 & 1 \\ 2 & -1 \\ -2 & 4 \end{pmatrix} \begin{pmatrix} 3 & 3 & 0 \\ 2 & 1 & 2 \end{pmatrix} \\
&= \frac{1}{9} \begin{pmatrix} 5 & 4 & 2 \\ 4 & 5 & -2 \\ 2 & -2 & 8 \end{pmatrix}
\end{aligned}$$

The results are identical in both cases. □

4.4. Fit a plane $z = a + bx + cy$ to four points $(1, 1, 3)$, $(0, 3, 6)$, $(2, 1, 5)$, and $(0, 0, 0)$. To do that

a) Find 4 equations with 3 unknowns a, b, c such that the plane passes through all 4 points (this system does not have to have a solution).

Answer.

$$\begin{aligned}
3 &= a + 1b + 1c \\
6 &= a + 0b + 3c \\
5 &= a + 2b + 1c \\
0 &= a + 0b + 0c
\end{aligned}$$

□

b) Find the least squares solution of the system.

Answer. Consider the system

$$\underbrace{\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 3 \\ 1 & 2 & 1 \\ 1 & 0 & 0 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} a \\ b \\ c \end{pmatrix}}_{\mathbf{x}} = \underbrace{\begin{pmatrix} 3 \\ 6 \\ 5 \\ 0 \end{pmatrix}}_{\mathbf{b}}$$

It follows from solving the normal equation that the least squares solution is

$$\begin{aligned}
\mathbf{x} &= (A^*A)^{-1}A^*\mathbf{b} \\
&= \left(\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 2 & 0 \\ 1 & 3 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 3 \\ 1 & 2 & 1 \\ 1 & 0 & 0 \end{pmatrix}\right)^{-1}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 2 & 0 \\ 1 & 3 & 1 & 0 \end{pmatrix}\begin{pmatrix} 3 \\ 6 \\ 5 \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} 4 & 3 & 5 \\ 3 & 5 & 3 \\ 4 & 3 & 8 \end{pmatrix}^{-1}\begin{pmatrix} 14 \\ 13 \\ 26 \end{pmatrix} \\
&= \frac{1}{33}\begin{pmatrix} 31 & -9 & -16 \\ -12 & 12 & 3 \\ -11 & 0 & 11 \end{pmatrix}\begin{pmatrix} 14 \\ 13 \\ 26 \end{pmatrix} \\
&= \frac{1}{50}\begin{pmatrix} -6 \\ 73 \\ 101 \end{pmatrix}
\end{aligned}$$

□

4.5. Minimal norm solution. Let an equation $A\mathbf{x} = \mathbf{b}$ have a solution, and let A have a non-trivial kernel (so the solution is not unique). Prove that

a) There exists a unique solution $\mathbf{x}_0$ of $A\mathbf{x} = \mathbf{b}$ minimizing the norm $\|\mathbf{x}\|$, i.e., that there exists a unique $\mathbf{x}_0$ such that $A\mathbf{x}_0 = \mathbf{b}$ and $\|\mathbf{x}_0\| \leq \|\mathbf{x}\|$ for any $\mathbf{x}$ satisfying $A\mathbf{x} = \mathbf{b}$.

Answer. Let $\mathbf{x}$ be a solution to $A\mathbf{x} = \mathbf{b}$, and choose $\mathbf{x}_0 = P_{(\ker A)^\perp}\mathbf{x}$. Then by Definition 3.1, $\mathbf{x} - \mathbf{x}_0 \in ((\ker A)^\perp)^\perp = \ker A$. Thus, since

$$\begin{aligned}
\mathbf{b} &= A\mathbf{x} \\
&= A(\mathbf{x} - \mathbf{x}_0 + \mathbf{x}_0) \\
&= A(\mathbf{x} - \mathbf{x}_0) + A\mathbf{x}_0 \\
&= A\mathbf{x}_0
\end{aligned}$$

we know that $\mathbf{x}_0$ is a solution of the equation $A\mathbf{x} = \mathbf{b}$. Now let $\mathbf{x}_h \in \ker A$. Then $\mathbf{x}_0 + \mathbf{x}_h$ is a solution to $A\mathbf{x} = \mathbf{b}$. Additionally, since $\mathbf{x}_0 \in (\ker A)^\perp$ by definition, we have by the Pythagorean theorem that

$$\|\mathbf{x}\|^2 = \|\mathbf{x}_0\|^2 + \|\mathbf{x}_h\|^2 \geq \|\mathbf{x}_0\|$$

so $\mathbf{x}_0$ is the unique minimal norm solution (any solution is of the form $\mathbf{x}_0 + \mathbf{x}_h$) because any other solution with a nontrivial $\mathbf{x}_h$ necessarily has a greater norm. □

b) $\mathbf{x}_0 = P_{(\ker A)^\perp}\mathbf{x}$ for any $\mathbf{x}$ satisfying $A\mathbf{x} = \mathbf{b}$.

Answer. See part (a). □

5.2. Find matrices of orthogonal projections onto all 4 fundamental subspaces of the matrix

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 2 & 4 & 3 \end{pmatrix}$$

Note that you only really need to compute 2 of the projections. If you pick an appropriate 2, the other 2 are easy to obtain from them (recall how the projections onto E and $E^\perp$ are related).

Answer. We can observe that $\dim \operatorname{range} A = 2$, so $\dim \ker A = 1$. Similarly, $\dim \operatorname{range} A^* = 2$, so $\dim \ker A^* = 1$. Since neither matrix is full rank, A^*A is singular so we cannot use the projection formula for any projection. Thus, we will use the Gram-Schmidt orthogonalization method to find the matrices of the projections onto $\ker A$ and $\ker A^*$ (since they're of lower dimension), and then we will invoke

$$\begin{aligned} I &= P_{\ker A} + P_{(\ker A)^\perp} & I &= P_{\ker A^*} + P_{(\ker A^*)^\perp} && \text{Exercise 5.3.13} \\ &= P_{\ker A} + P_{\operatorname{range} A^*} & &= P_{\ker A^*} + P_{\operatorname{range} A} && \text{Theorem 5.1} \end{aligned}$$

to find the other two. Let's begin.

First off, we have that

$$\ker A = \left\{ \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix} \right\} \qquad \ker A^* = \left\{ \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} \right\}$$

Thus, we have from the Gram-Schmidt orthogonalization formula in the one-dimensional case that

$$\begin{aligned} P_{\ker A} &= \frac{1}{(-1)^2 + (-1)^2 + 2^2} \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix} \begin{pmatrix} -1 & -1 & 2 \end{pmatrix} & P_{\ker A^*} &= \frac{1}{(-1)^2 + (-1)^2 + 1^2} \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} \begin{pmatrix} -1 & -1 & 1 \end{pmatrix} \\ &= \frac{1}{6} \begin{pmatrix} 1 & 1 & -2 \\ 1 & 1 & -2 \\ -2 & -2 & 4 \end{pmatrix} & &= \frac{1}{3} \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix} \end{aligned}$$

It follows that

$$P_{\operatorname{range} A^*} = \frac{1}{6} \begin{pmatrix} 5 & -1 & 2 \\ -1 & 5 & 2 \\ 2 & 2 & 2 \end{pmatrix} \qquad P_{\operatorname{range} A} = \frac{1}{3} \begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

$\square$

5.3. Let A be an $m \times n$ matrix. Show that $\ker A = \ker(A^*A)$. (Hint: To do this, you need to prove 2 inclusions, namely $\ker(A^*A) \subset \ker A$ and $\ker A \subset \ker(A^*A)$. One of the inclusions is trivial, and for the other one, use the fact that $\|A\mathbf{x}\|^2 = (A\mathbf{x}, A\mathbf{x}) = (A^*A\mathbf{x}, \mathbf{x})$.)

Answer. Suppose $\mathbf{x} \in \ker A$. Then $A\mathbf{x} = \mathbf{0}$. It follows that $A^*A\mathbf{x} = A^*\mathbf{0} = \mathbf{0}$, so $\mathbf{x} \in \ker(A^*A)$, as desired.

Now suppose that $\mathbf{x} \in \ker(A^*A)$. Then $A^*A\mathbf{x} = \mathbf{0}$. It follows that

$$0 = (\mathbf{0}, \mathbf{x}) = (A^*A\mathbf{x}, \mathbf{x}) = (A\mathbf{x}, A\mathbf{x}) = \|A\mathbf{x}\|^2$$

But this implies that $A\mathbf{x} = 0$, so $\mathbf{x} \in \ker A$, as desired. $\square$

5.4. Use the equality $\ker A = \ker(A^*A)$ to prove that

a) $\operatorname{rank} A = \operatorname{rank}(A^*A)$.

Answer. Let A be $m \times n$. Then by consecutive applications of the rank theorem,

$$\dim \ker A + \operatorname{rank} A = n \qquad \dim \ker(A^*A) + \operatorname{rank}(A^*A) = n$$

It follows since $\ker A = \ker(A^*A)$ that

$$\operatorname{rank} A = n - \dim \ker A = n - \dim \ker(A^*A) = \operatorname{rank}(A^*A)$$

as desired. $\square$

b) If $A\mathbf{x} = \mathbf{0}$ has only the trivial solution, A is left invertible. (Hint: You can just write a formula for the left inverse.)

Answer. If $A\mathbf{x} = \mathbf{0}$ has only the trivial solution, then A is full rank. It follows by the above that A^*A is full rank. Additionally, however, A^*A is square, so A^*A is invertible. Therefore, there exists a matrix $(A^*A)^{-1}$ such that $(A^*A)^{-1}A^*A = I$, but this implies that the left inverse of A is just $(A^*A)^{-1}A^*$. □

5.6. Let a matrix P be self-adjoint ($P^* = P$) and let $P^2 = P$. Show that P is the matrix of an orthogonal projection. (Hint: Consider the decomposition $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$, $\mathbf{x}_1 \in \operatorname{range} P$ and $\mathbf{x}_2 \perp \operatorname{range} P$, and show that $P\mathbf{x}_1 = \mathbf{x}_1$, $P\mathbf{x}_2 = \mathbf{0}$. For one of the equalities, you will need self-adjointness; for the other one, the property $P^2 = P$.)

Answer. We will prove that $P : V \to V$ self-adjoint and satisfying $P^2 = P$ is the matrix of the orthogonal projection onto $\operatorname{range} P$. Let $\mathbf{x} \in V$. Then since $V = \operatorname{range} P \oplus (\operatorname{range} P)^\perp$, we may let $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$ where $\mathbf{x}_1 \in \operatorname{range} P$ and $\mathbf{x}_2 \in (\operatorname{range} P)^\perp$. Taking the hint, we now prove that $P\mathbf{x}_1 = \mathbf{x}_1$ and $P\mathbf{x}_2 = \mathbf{0}$. Let's begin.

Since $\mathbf{x}_1 \in \operatorname{range} P$, we have that $\mathbf{x}_1 = P\mathbf{y}$ for some $\mathbf{y} \in V$. But since $P^2 = P$, we have that

$$\mathbf{x}_1 = P\mathbf{y} = P^2\mathbf{y} = P(P\mathbf{y}) = P\mathbf{x}_1$$

as desired.

On the other hand, $\mathbf{x}_2 \in (\operatorname{range} P)^\perp = (\operatorname{range} P^*)^\perp = \ker P$ by Theorem 5.1, so naturally $P\mathbf{x}_2 = \mathbf{0}$.

Therefore,

$$P\mathbf{x} = P(\mathbf{x}_1 + \mathbf{x}_2) = P\mathbf{x}_1 + P\mathbf{x}_2 = \mathbf{x}_1$$

as desired for an orthogonal projection onto $\operatorname{range} P$. □

6.1. Orthogonally diagonalize the following matrices

$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \qquad \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 2 & 2 \\ 2 & 0 & 2 \\ 2 & 2 & 0 \end{pmatrix}$$

i.e., for each matrix A, find a unitary matrix U and a diagonal matrix D such that $A = UDU^*$.

Answer. The leftmost matrix has eigenvalues $\lambda = 3, -1$ and corresponding orthonormal eigenvectors

$$x_1 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \qquad x_2 = \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$$

Therefore,

$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$$

The middle matrix has eigenvalues $\lambda = i, -i$ and corresponding orthonormal eigenvectors

$$x_1 = \begin{pmatrix} i/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \qquad x_2 = \begin{pmatrix} -i/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$$

Therefore,

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} i/\sqrt{2} & -i/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \begin{pmatrix} -i/\sqrt{2} & 1/\sqrt{2} \\ i/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$$

The right matrix has eigenvalues $\lambda = -2, 4$ (-2 having multiplicity 2) and corresponding orthonormal eigenvectors

$$x_1 = \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix} \qquad x_2 = \begin{pmatrix} -1/\sqrt{6} \\ -1/\sqrt{6} \\ \sqrt{6}/3 \end{pmatrix} \qquad x_3 = \begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix}$$

Therefore,

$$\begin{pmatrix} 0 & 2 & 2 \\ 2 & 0 & 2 \\ 2 & 2 & 0 \end{pmatrix} = \begin{pmatrix} -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & \sqrt{6}/3 & 1/\sqrt{3} \end{pmatrix} \begin{pmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} -1/\sqrt{2} & 1/\sqrt{2} & 0 \\ -1/\sqrt{6} & -1/\sqrt{6} & \sqrt{6}/3 \\ 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \end{pmatrix}$$

□

6.2. True or false: A matrix is unitarily equivalent to a diagonal one if and only if it has an orthogonal basis of eigenvectors.

Answer. True.

See Proposition 6.5. □

6.5. Let $U : X \to X$ be a linear transformation on a finite-dimensional inner product space. True or false:

a) If $\|U\mathbf{x}\| = \|\mathbf{x}\|$ for all $\mathbf{x} \in X$, then U is unitary.

Answer. True.

By Proposition 6.3 — the fact that $\|U\mathbf{x}\| = \|\mathbf{x}\|$ for all $\mathbf{x} \in X$ makes U an isometry; clearly $\dim X = \dim X$. □

b) If $\|U\mathbf{e}_k\| = \|\mathbf{e}_k\|$ for each $k = 1, \dots, n$ for some orthonormal basis $\mathbf{e}_1, \dots, \mathbf{e}_n$, then U is unitary.

Answer. False.

Let $U\mathbf{e}_k = \mathbf{e}_1$ for all $k = 1, \dots, n$. Then the identity holds. However, $U\mathbf{e}_1, \dots, U\mathbf{e}_n$ is not an orthonormal basis of X, so U cannot be unitary. □

6.6. Let A and B be unitarily equivalent $n \times n$ matrices.

a) Prove that $\operatorname{tr}(A^*A) = \operatorname{tr}(B^*B)$.

Answer. Let $A = UBU^*$. Then

$$\begin{aligned} \operatorname{tr}(A^*A) &= \operatorname{tr}((UBU^*)^*(UBU^*)) \\ &= \operatorname{tr}((U^*)^*B^*U^*UBU^*) \\ &= \operatorname{tr}(UB^*IBU^*) \\ &= \operatorname{tr}(B^*BUU^*) \\ &= \operatorname{tr}(B^*BI) \\ &= \operatorname{tr}(B^*B) \end{aligned}$$

as desired. □

b) Use (a) to prove that

$$\sum_{j,k=1}^{n} |A_{j,k}|^2 = \sum_{j,k=1}^{n} |B_{j,k}|^2$$

Answer. By the definition of the adjoint, we have that

$$|A_{j,k}|^2 = \overline{A_{j,k}} A_{j,k} = A^*_{k,j} A_{j,k}$$

This allows us to show that

$$(A^*A)_{k,k} = \sum_{j=1}^{n} A^*_{k,j} A_{j,k} = \sum_{j=1}^{n} |A_{j,k}|^2$$

Therefore, we have that

$$\begin{aligned}
\sum_{j,k=1}^{n} |A_{j,k}|^2 &= \sum_{k=1}^{n} \sum_{j=1}^{n} |A_{j,k}|^2 \\
&= \sum_{k=1}^{n} (A^*A)_{k,k} \\
&= \operatorname{tr}(A^*A) \\
&= \operatorname{tr}(B^*B) \\
&= \sum_{k=1}^{n} (B^*B)_{k,k} \\
&= \sum_{k=1}^{n} \sum_{j=1}^{n} |B_{j,k}|^2 \\
&= \sum_{j,k=1}^{n} |A_{j,k}|^2
\end{aligned}$$

as desired □

c) Use (b) to prove that the matrices

$$\begin{pmatrix} 1 & 2 \\ 2 & i \end{pmatrix} \qquad \begin{pmatrix} i & 4 \\ 1 & 1 \end{pmatrix}$$

are not unitarily equivalent.

Answer. We have that

$$\begin{aligned}
\sum_{j,k=1}^{2} |A_{j,k}|^2 &= |1|^2 + |2|^2 + |2|^2 + |i|^2 \\
&= 10
\end{aligned}
\qquad
\begin{aligned}
\sum_{j,k=1}^{n} |B_{j,k}|^2 &= |i|^2 + |4|^2 + |1|^2 + |1|^2 \\
&= 19
\end{aligned}$$

Therefore, by part (b), the above matrices are not unitarily equivalent. □

6.7. Which of the following pairs of matrices are unitarily equivalent? (Hint: It is easy to eliminate matrices that are not unitarily equivalent: Remember that unitarily equivalent matrices are similar, and recall that the trace, determinant, and eigenvalues of similar matrices coincide. Also, the previous problem helps in eliminating non-unitarily equivalent matrices. Finally, a matrix is unitarily equivalent to a diagonal one if and only if it has an orthogonal basis of eigenvectors.)

a)

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Answer. Not unitarily equivalent.
The traces of the two matrices above differ. □

b)

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1/2 \\ 1/2 & 0 \end{pmatrix}$$

Answer. Not unitarily equivalent.
The determinants of the two matrices above differ. □

c)

$$\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Answer. Not unitarily equivalent.
The eigenvalues of the two matrices above differ. □

d)

$$\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 \\ 0 & -i & 0 \\ 0 & 0 & i \end{pmatrix}$$

Answer. Unitarily equivalent.
We have that

$$\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -i/\sqrt{2} & i/\sqrt{2} \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -i & 0 \\ 0 & 0 & i \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ i/\sqrt{2} & 1/\sqrt{2} & 0 \\ -i/\sqrt{2} & 1/\sqrt{2} & 0 \end{pmatrix}$$

□

e)

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

Answer. Not unitarily equivalent. The sums from Problem 6.6b are not equal. □

6.8. Let U be a 2×2 orthogonal matrix with $\det U = 1$. Prove that U is a rotation matrix.

Answer. Since U is orthogonal and real, $U^T = U^* = U^{-1}$. It follows that

$$\begin{aligned} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} a & c \\ b & d \end{pmatrix} \\ &= \begin{pmatrix} a^2 + b^2 & ac + bd \\ ca + db & c^2 + d^2 \end{pmatrix} \end{aligned}$$

If $a^2 + b^2 = 1$, then we must have $a = \cos\gamma$, $b = -\sin\gamma$ for some γ. Similarly, we must have $c = \cos\theta$ and $d = \sin\theta$. We have now expressed our four variables in terms of two. To get it down to one, we apply the determinant condition:

$$\begin{aligned} 1 &= \det U \\ &= ad - bc \\ &= \cos\gamma\sin\theta + \sin\gamma\cos\theta \\ &= \sin(\theta + \gamma) \\ \theta + \gamma &= \frac{\pi}{2} + 2\pi n \\ \theta &= \frac{\pi}{2} - \gamma + 2\pi n \end{aligned}$$

where $n \in \mathbb{Z}$. It follows that

$$c = \cos\theta = \cos\left(\frac{\pi}{2} - \gamma + 2\pi n\right) = \sin(\gamma + 2\pi n) = \sin\gamma$$

$$d = \sin\theta = \sin\left(\frac{\pi}{2} - \gamma + 2\pi n\right) = \cos(\gamma + 2\pi n) = \cos\gamma$$

Therefore,

$$U = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \cos\gamma & -\sin\gamma \\ \sin\gamma & \cos\gamma \end{pmatrix} = R_\gamma$$

as desired. □

6.9. Let U be a 3×3 orthogonal matrix with $\det U = 1$. Prove that

a) 1 is an eigenvalue of U.

Answer. Let λ be an eigenvalue of U with corresponding eigenvector $\mathbf{x}$. Then

$$\begin{aligned} |\lambda|\|\mathbf{x}\| = \|\lambda\mathbf{x}\| &= \|U\mathbf{x}\| = \|\mathbf{x}\| \\ |\lambda| &= 1 \end{aligned}$$

It follows that

$$\lambda_1 = \mathrm{e}^{ix_1} \qquad \lambda_2 = \mathrm{e}^{ix_2} \qquad \lambda_3 = \mathrm{e}^{ix_3}$$

We know that if any eigenvalue is complex, its complex conjugate must also be an eigenvalue. Thus, WLOG let $x_2 = -x_1$ so

$$\lambda_1 = \mathrm{e}^{ix_1} \qquad \lambda_2 = \mathrm{e}^{-ix_1} \qquad \lambda_3 = \mathrm{e}^{ix_3}$$

Then

$$1 = \det U = \lambda_1\lambda_2\lambda_3 = \mathrm{e}^{ix_1}\mathrm{e}^{-ix_1}\mathrm{e}^{ix_3} = \mathrm{e}^{ix_3} = \lambda_3$$

as desired. □

b) If $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ is an orthonormal basis, such that $U\mathbf{v}_1 = \mathbf{v}_1$ (remember that 1 is an eigenvalue), then in this basis, the matrix of U is

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}$$

where α is some angle. (Hint: Show that since $\mathbf{v}_1$ is an eigenvector of U, all entries below 1 must be zero, and since $\mathbf{v}_1$ is also an eigenvector of U^* [why?], all entries right of 1 must also be zero. Then show that the lower right 2×2 matrix is an orthogonal one with determinant 1, and use the previous problem.)

Answer. Since

$$U\mathbf{v}_1 = \mathbf{v}_1 = 1\mathbf{v}_1 + 0\mathbf{v}_2 + 0\mathbf{v}_3$$

we must have that the first column of U with respect to $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ is

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

Additionally, since

$$\mathbf{v}_1 = I\mathbf{v}_1 = U^*U\mathbf{v}_1 = U^*\mathbf{v}_1$$

we know that $\mathbf{v}_1$ is an eigenvector of U^* with corresponding eigenvalue 1. Thus, by the above, the first column of U^* is also

$$\begin{pmatrix}1\\0\\0\end{pmatrix}$$

meaning that the first row of U is

$$\begin{pmatrix}1 & 0 & 0\end{pmatrix}$$

Thus, if we block diagonalize U, we have that

$$\begin{pmatrix}1 & 0\\0 & I\end{pmatrix} = \begin{pmatrix}1 & 0\\0 & L\end{pmatrix}^* \begin{pmatrix}1 & 0\\0 & L\end{pmatrix} = \begin{pmatrix}1 & 0\\0 & L^*\end{pmatrix} \begin{pmatrix}1 & 0\\0 & L\end{pmatrix}$$

so $L^*L = I$. Thus, by Lemma 6.2, L is an isometry. Additionally, by Proposition 6.3, L is unitary. Furthermore, since U is orthogonal, i.e., has all real values, L must have all real values and must be orthogonal, too. Lastly, since $\det U = 1$, we have by the method of cofactor expansion that

$$1 = \det U = 1 \cdot \det L = \det L$$

Therefore, since L is a 2×2 orthogonal matrix with $\det L = 1$, Exercise 6.8 implies that $U = R_\alpha$ for some α, as desired. ☐

8.1. Prove the following formula.

$$(\mathbf{x}, \mathbf{y})_{\mathbb{R}} = \operatorname{Re}(\mathbf{x}, \mathbf{y})_{\mathbb{C}}$$

Namely, show that if

$$\mathbf{x} = \begin{pmatrix}z_1\\ \vdots \\ z_n\end{pmatrix} \qquad \mathbf{y} = \begin{pmatrix}w_1\\ \vdots \\ w_n\end{pmatrix}$$

where $z_k = x_k + iy_k$, $w_k = u_k + iv_k$, $x_k, y_k, u_k, v_k \in \mathbb{R}$, then

$$\operatorname{Re}\left(\sum_{k=1}^n z_k \bar{w}_k\right) = \sum_{k=1}^n x_k u_k + \sum_{k=1}^n y_k v_k$$

Answer. We have that

$$\begin{aligned}\operatorname{Re}\left(\sum_{k=1}^n z_k \bar{w}_k\right) &= \operatorname{Re}\left(\sum_{k=1}^n (x_k + iy_k)(u_k - iv_k)\right)\\ &= \operatorname{Re}\left(\sum_{k=1}^n [x_k u_k + y_k v_k + i(y_k u_k - x_k v_k)]\right)\\ &= \sum_{k=1}^n [x_k u_k + y_k v_k]\\ &= \sum_{k=1}^n x_k u_k + \sum_{k=1}^n y_k v_k\end{aligned}$$

as desired. ☐

8.4. Show that if U is an orthogonal transformation satisfying $U^2 = -I$, then $U^* = -U$.

Answer. Suppose $U^2 = -I$. This combined with the fact that $U^*U = I$ implies that

$$\begin{aligned}
U^*U + UU &= 0 \\
(U^* + U)U &= 0 \\
(U^* + U)UU^{-1} &= 0U^{-1} \\
U^* + U &= 0 \\
U^* &= -U
\end{aligned}$$

as desired. □

4 Inner Product Phenomena and Intro to Bilinear Forms

From Treil (2017).

Chapter 6

10/25: **1.1.** Use the upper-triangular representation of an operator to give an alternative proof of the fact that the determinant is the product and the trace is the sum of the eigenvalues counting multiplicities.

Answer. Let $A : V \to V$ be an operator. Then by Theorem 6.1.1, there exists a basis of V such that the matrix of A with respect to this basis is upper triangular. Since this matrix is upper triangular, the eigenvalues of A are exactly its diagonal entries. This combined with the fact that the determinant of an upper triangular matrix is the product of its diagonal entries proves that the determinant of A is the product of its eigenvalues. Similarly, the trace of A as the sum of the diagonal entries of A must be the sum of the eigenvalues of A, as desired. □

2.1. True or false:

a) Every unitary operator $U : X \to X$ is normal.

Answer. True.
Let $U : X \to X$ be unitary. Then

$$U^*U = I = UU^*$$

as desired. □

b) A matrix is unitary if and only if it is invertible.

Answer. False.
Consider the matrix

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

A is invertible with inverse

$$A^{-1} = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}$$

but A is not unitary since A is not an isometry:

$$\|A\mathbf{x}\| = \|2\mathbf{x}\| = 2\|\mathbf{x}\| \neq \|\mathbf{x}\|$$

for any $\mathbf{x} \in \mathbb{F}^2$. □

c) If two matrices are unitarily equivalent, then they are also similar.

Answer. True.
Suppose that $A = UBU^*$. Then since $U^* = U^{-1}$, $A = UBU^{-1}$, so A, B are similar. □

d) The sum of self-adjoint operators is self-adjoint.

Answer. True.
If $A = A^*$ and $B = B^*$, then

$$(A + B)^* = A^* + B^* = A + B$$

as desired. □

e) The adjoint of a unitary operator is unitary.

Answer. True.
See property 2 of unitary operators (Treil, 2017, p. 148). □

f) The adjoint of a normal operator is normal.

Answer. True.
Let N be normal. Then $N^*N = NN^*$. This combined with the fact that $N = (N^*)^*$ implies that

$$(N^*)^*N^* = NN^* = N^*N = N^*(N^*)^*$$

as desired. □

g) If all eigenvalues of a linear operator are 1, then the operator must be unitary or orthogonal.

Answer. False.
Consider the matrix

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

Clearly all eigenvalues of this matrix are 1. However,

$$A^*A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \neq I$$

so A is not unitary. □

h) If all eigenvalues of a normal operator are 1, then the operator is the identity.

Answer. True.
Suppose N is a normal operator with all eigenvalues equal to 1. Then by Theorem 6.2.4, $N = UDU^*$ where $D = I$ (because of the condition on the eigenvalues). It follows that

$$N = UIU^* = UU^* = I$$

as desired. □

i) A linear operator may preserve norm but not the inner product.

Answer. False.
Suppose U is a linear operator that preserves norm. Then U is an isometry. It follows by Theorem 5.6.1 that U preserves the inner product. □

2.2. True or false (justify your conclusion): The sum of normal operators is normal.

Answer. False.
Let

$$N = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \qquad M = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

We know that N, M are normal since

$$\begin{aligned} NN^* &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \\ &= N^*N \end{aligned} \qquad \begin{aligned} MM^* &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ &= M^*M \end{aligned}$$

Then we have

$$\begin{aligned}(N+M)(N+M)^* &= \begin{pmatrix}0 & 2\\ 0 & 0\end{pmatrix}\begin{pmatrix}0 & 0\\ 2 & 0\end{pmatrix}\\ &= \begin{pmatrix}4 & 0\\ 0 & 0\end{pmatrix}\\ &\neq \begin{pmatrix}0 & 0\\ 0 & 4\end{pmatrix}\\ &= \begin{pmatrix}0 & 0\\ 2 & 0\end{pmatrix}\begin{pmatrix}0 & 2\\ 0 & 0\end{pmatrix}\\ &= (N+M)^*(N+M)\end{aligned}$$

□

2.3. Show that an operator that is unitarily equivalent to a diagonal one is normal.

Answer. Let $A = UDU^*$. Then

$$\begin{aligned}NN^* &= (UDU^*)(UDU^*)^* & N^*N &= (UDU^*)^*(UDU^*)\\ &= (UDU^*)(UD^*U^*) & &= (UD^*U^*)(UDU^*)\\ &= UDD^*U^* & &= UD^*DU^*\end{aligned}$$

Additionally, we have that $D^*D = DD^*$ (Treil, 2017, p. 167), completing the proof. □

2.5. True or false (justify): Any self-adjoint matrix has a self-adjoint square root.

Answer. False.
Consider the trivially self-adjoint matrix

$$\begin{pmatrix}-1\end{pmatrix}$$

The square roots of this matrix are $\begin{pmatrix}i\end{pmatrix}$ and $\begin{pmatrix}-i\end{pmatrix}$, neither of which is self-adjoint. □

2.6. Orthogonally diagonalize the matrix

$$A = \begin{pmatrix}7 & 2\\ 2 & 4\end{pmatrix}$$

i.e., represent it as $A = UDU^*$, where D is diagonal and U is unitary. Additionally, among all square roots of A, i.e., among all matrices B such that $B^2 = A$, find one that has positive eigenvalues. You can leave B as a product.

Answer. From the characteristic polynomial, we find that $\lambda_1 = 8$ and $\lambda_2 = 3$. It follows by inspection that the corresponding eigenvectors are

$$\mathbf{x}_1 = \begin{pmatrix}2\\ 1\end{pmatrix} \qquad \mathbf{x}_2 = \begin{pmatrix}-1\\ 2\end{pmatrix}$$

These vectors are already orthogonal, so we need only normalize them to get

$$U = \begin{pmatrix}2/\sqrt{5} & -1/\sqrt{5}\\ 1/\sqrt{5} & 2/\sqrt{5}\end{pmatrix}$$

Therefore, we have that

$$A = \begin{pmatrix}2/\sqrt{5} & -1/\sqrt{5}\\ 1/\sqrt{5} & 2/\sqrt{5}\end{pmatrix}\begin{pmatrix}8 & 0\\ 0 & 3\end{pmatrix}\begin{pmatrix}2/\sqrt{5} & 1/\sqrt{5}\\ -1/\sqrt{5} & 2/\sqrt{5}\end{pmatrix}$$

From here, we can easily let

$$B = \begin{pmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix} \begin{pmatrix} 2\sqrt{2} & 0 \\ 0 & \sqrt{3} \end{pmatrix} \begin{pmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix}$$

□

2.7. True or false (justify your conclusions):

a) A product of two self-adjoint matrices is self-adjoint.

Answer. False.
Let

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$$

Clearly $A = A^*$ and $B = B^*$. However,

$$AB = \begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix} = (AB)^*$$

□

b) If A is self-adjoint, then A^k is self-adjoint.

Answer. True.
Suppose $A = A^*$. Then

$$(A^k)^* = (\underbrace{A \cdots A}_{k \text{ times}})^* = \underbrace{A^* \cdots A^*}_{k \text{ times}} = \underbrace{A \cdots A}_{k \text{ times}} = A^k$$

as desired. □

2.8. Let A be an $m \times n$ matrix. Prove that

a) A^*A is self-adjoint.

Answer. We have that

$$(A^*A)^* = A^*(A^*)^* = A^*A$$

as desired. □

b) All eigenvalues of A^*A are nonnegative.

Answer. Let λ be an eigenvalue of A^*A with corresponding nonzero eigenvector $\mathbf{x}$. Then

$$0 \leq (A\mathbf{x}, A\mathbf{x}) = (A^*A\mathbf{x}, \mathbf{x}) = (\lambda\mathbf{x}, \mathbf{x}) = \lambda(\mathbf{x}, \mathbf{x}) = \lambda\|\mathbf{x}\|^2$$

$$\frac{0}{\|\mathbf{x}\|^2} = 0 \leq \lambda$$

as desired. □

c) $A^*A + I$ is invertible.

Answer. To show that $A^*A + I$ is invertible, it will suffice to show that $\ker(A^*A + I) = \{\mathbf{0}\}$. One inclusion is obvious. However, for the other one, suppose $(A^*A + I)\mathbf{x} = \mathbf{0}$. Then

$$\begin{aligned} 0 &= (\mathbf{0}, \mathbf{x}) \\ &= ((A^*A + I)\mathbf{x}, \mathbf{x}) \\ &= (A^*A\mathbf{x} + \mathbf{x}, \mathbf{x}) \\ &= (A^*A\mathbf{x}, \mathbf{x}) + (\mathbf{x}, \mathbf{x}) \\ &= (A\mathbf{x}, A\mathbf{x}) + (\mathbf{x}, \mathbf{x}) \\ &= \|A\mathbf{x}\|^2 + \|\mathbf{x}\|^2 \end{aligned}$$

Therefore, $\|\mathbf{x}\| = 0$, so $\mathbf{x} = \mathbf{0}$, so $\mathbf{x} \in \{\mathbf{0}\}$, as desired. □

2.10. Orthogonally diagonalize the rotation matrix

$$R_\alpha = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

where α is not a multiple of π. Note that you will get complex eigenvalues in this case.

Answer. We have from Problem 4.1.3 that $\lambda_1 = \mathrm{e}^{i\alpha}$ and $\lambda_2 = \mathrm{e}^{-i\alpha}$, and that

$$\mathbf{x}_1 = \begin{pmatrix} 1 \\ -i \end{pmatrix} \qquad \mathbf{x}_2 = \begin{pmatrix} 1 \\ i \end{pmatrix}$$

where $\mathbf{x}_1, \mathbf{x}_2$ are already orthogonal. Thus, normalizing gives us

$$R_\alpha = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -i/\sqrt{2} & i/\sqrt{2} \end{pmatrix} \begin{pmatrix} \mathrm{e}^{i\alpha} & 0 \\ 0 & \mathrm{e}^{-i\alpha} \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & i/\sqrt{2} \\ 1/\sqrt{2} & -i/\sqrt{2} \end{pmatrix}$$

□

2.13. Prove that a normal operator with unimodular eigenvalues (i.e., with all eigenvalues satisfying $|\lambda_k| = 1$) is unitary. (Hint: Consider diagonalization.)

Answer. Let N be normal with unimodular eigenvalues. To prove that N is unitary, it will suffice to show that $NN^* = I$. First off, we have by Theorem 6.2.4 that $N = UDU^*$ where U is unitary and D is diagonal. Thus,

$$NN^* = UDU^*(UDU^*)^* = UDU^*UD^*U^* = UDD^*U = UIU^* = I$$

as desired. Note that $DD^* = I$ since each value along the diagonal of DD^* has $d_{jj}\bar{d}_{jj} = |d_{jj}|^2 = 1$. □

2.14. Prove that a normal operator with real eigenvalues is self-adjoint.

Answer. Let N be normal with all real eigenvalues. By Theorem 6.2.4, $N = UDU^*$ where D is real. Then

$$N^* = (UDU^*)^* = UD^*U^* = UDU^* = N$$

as desired. □

2.15. Show by example that the conclusion of Theorem 2.2 fails for *complex* symmetric matrices. Namely,

a) Construct a (diagonalizable) 2×2 complex symmetric matrix not admitting an orthogonal basis of eigenvectors.

Answer. Suppose A is our final matrix. We will apply the constraints sequentially to narrow down possible values of A and then pick one. Let's begin.
Since A is diagonalizable, $A = SDS^{-1}$ where D is a diagonal matrix and S is a matrix of eigenvectors of A. Since A is symmetric, $A = A^T$. It follows from these two conditions that

$$\begin{aligned} SDS^{-1} &= (SDS^{-1})^T \\ SDS^{-1} &= (S^T)^{-1}D^TS^T \\ S^TSD &= DS^TS \end{aligned}$$

Since $(S^TS)^T = S^T(S^T)^T = S^TS$ (so S^TS is symmetric), D is diagonal, and both are 2×2, we can represent them as

$$S^TS = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \qquad D = \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix}$$

for some $a, b, c, d_1, d_2 \in \mathbb{C}$. Thus, the above condition implies that

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix} = \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix} \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$
$$\begin{pmatrix} ad_1 & bd_2 \\ bd_1 & cd_2 \end{pmatrix} = \begin{pmatrix} ad_1 & bd_1 \\ bd_2 & cd_2 \end{pmatrix}$$

i.e., that $bd_1 = bd_2$. It follows that either $d_1 = d_2$, or $b = 0$. Since we would like the freedom to choose distinct values, we will choose a solution for which $b = 0$. The overall conclusion is that $S^T S$ is diagonal, which implies that $\mathbf{x}_2^T \mathbf{x}_1 = 0$.

We now invoke the last given condition: that the eigenvectors $\mathbf{x}_1, \mathbf{x}_2$ are not orthogonal, i.e., $\mathbf{x}_2^* \mathbf{x}_1 \neq 0$.

To summarize, our final matrix is of the form

$$A = \begin{pmatrix} \mathbf{x}_1 & \mathbf{x}_2 \end{pmatrix} \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix} \begin{pmatrix} \mathbf{x}_1 & \mathbf{x}_2 \end{pmatrix}^{-1}$$

We need to choose $\mathbf{x}_1, \mathbf{x}_2$ such that $\mathbf{x}_2^T \mathbf{x}_1 = 0$, $\mathbf{x}_2^* \mathbf{x}_1 \neq 0$, and (of course) $\mathbf{x}_1, \mathbf{x}_2$ are linearly independent. And we need to choose d_1, d_2 such that the final matrix is complex (and it'd be nice if they put A in an easily readable form). For the eigenvectors, we can find the following two satisfactory eigenvectors by inspection.

$$\mathbf{x}_1 = \begin{pmatrix} 2 \\ i \end{pmatrix} \qquad \mathbf{x}_2 = \begin{pmatrix} 1 \\ 2i \end{pmatrix}$$

For the corresponding eigenvalues, it is easy to see that 3 and -3 nicely fit the bill, yielding

$$A = \begin{pmatrix} 2 & 1 \\ i & 2i \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & -3 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ i & 2i \end{pmatrix}^{-1} = \begin{pmatrix} 5 & 4i \\ 4i & -5 \end{pmatrix}$$

as our final diagonalizable 2×2 complex symmetric matrix that does not admit an orthogonal basis of eigenvectors. □

b) Construct a 2×2 complex symmetric matrix which cannot be diagonalized.

Answer. We have

$$\begin{pmatrix} 0 & 1 \\ 1 & i \end{pmatrix}$$

as a complex symmetric matrix that cannot be diagonalized. □

3.1. Show that the number of nonzero singular values of a matrix A coincides with its rank.

Answer. By Problem 5.5.4a, $\operatorname{rank} A = \operatorname{rank} A^*A$. Additionally, since A^*A is self-adjoint by Problem 6.2.8a, we have by Theorem 6.2.1 that A^*A is similar to a diagonal matrix D. Since similar matrices have the same rank, $\operatorname{rank}(A^*A) = \operatorname{rank}(D)$. But $\operatorname{rank}(D)$ is just the number of nonzero entries on the diagonal, i.e., the number of eigenvalues of A^*A. Therefore, since the singular values of A are the square roots of the eigenvalues of A^*A, the number of nonzero singular values of A equals the number of nonzero eigenvalues of A^*A. □

3.2. Find Schmidt decompositions $A = \sum_{k=1}^{r} s_k \mathbf{w}_k \mathbf{v}_k^*$ for the following matrices A.

$$\begin{pmatrix} 2 & 3 \\ 0 & 2 \end{pmatrix} \qquad \begin{pmatrix} 7 & 1 \\ 0 & 0 \\ 5 & 5 \end{pmatrix} \qquad \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ -1 & 1 \end{pmatrix}$$

Answer. Left matrix: We have

$$A^*A = \begin{pmatrix} 2 & 0 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 4 & 6 \\ 6 & 13 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{5} & -2/\sqrt{5} \\ 2/\sqrt{5} & 1/\sqrt{5} \end{pmatrix} \begin{pmatrix} 16 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ -2/\sqrt{5} & 1/\sqrt{5} \end{pmatrix}$$

so that $\sigma_1 = 4$ and $\sigma_2 = 1$, and

$$\mathbf{v}_1 = \begin{pmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix} \qquad \mathbf{v}_2 = \begin{pmatrix} -2/\sqrt{5} \\ 1/\sqrt{5} \end{pmatrix}$$

Then

$$\begin{aligned} \mathbf{w}_1 &= \frac{1}{\sigma_1} A\mathbf{v}_1 \\ &= \frac{1}{4} \begin{pmatrix} 2 & 3 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix} \\ &= \begin{pmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{pmatrix} \end{aligned} \qquad \begin{aligned} \mathbf{w}_2 &= \frac{1}{\sigma_2} A\mathbf{v}_2 \\ &= \frac{1}{1} \begin{pmatrix} 2 & 3 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} -2/\sqrt{5} \\ 1/\sqrt{5} \end{pmatrix} \\ &= \begin{pmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix} \end{aligned}$$

Thus, we have as a Schmidt decomposition

$$\begin{pmatrix} 2 & 3 \\ 0 & 2 \end{pmatrix} = 4 \begin{pmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{pmatrix} \begin{pmatrix} 1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix} + 1 \begin{pmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix} \begin{pmatrix} -2/\sqrt{5} & 1/\sqrt{5} \end{pmatrix}$$

Middle matrix: We have

$$A^*A = \begin{pmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix} \begin{pmatrix} 90 & 0 \\ 0 & 10 \end{pmatrix} \begin{pmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix}$$

so that $\sigma_1 = 3\sqrt{10}$ and $\sigma_2 = \sqrt{10}$, and

$$\mathbf{v}_1 = \begin{pmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{pmatrix} \qquad \mathbf{v}_2 = \begin{pmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix}$$

Then

$$\mathbf{w}_1 = \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{pmatrix} \qquad \mathbf{w}_2 = \begin{pmatrix} -1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{pmatrix}$$

Thus, we have as a Schmidt decomposition

$$\begin{pmatrix} 7 & 1 \\ 0 & 0 \\ 5 & 5 \end{pmatrix} = 3\sqrt{10} \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 2/\sqrt{5} & 1/\sqrt{5} \end{pmatrix} + \sqrt{10} \begin{pmatrix} -1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} -1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix}$$

Right matrix: We have

$$A^*A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

so that $\sigma_1 = \sqrt{2}$ and $\sigma_2 = \sqrt{3}$, and

$$\mathbf{v}_1 = \mathbf{e}_1 \qquad \mathbf{v}_2 = \mathbf{e}_2$$

Then

$$\mathbf{w}_1 = \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{pmatrix} \qquad \mathbf{w}_2 = \begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix}$$

Thus, we have as a Schmidt decomposition

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \\ -1 & 1 \end{pmatrix} = \sqrt{2} \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 1 & 0 \end{pmatrix} + \sqrt{3} \begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix} \begin{pmatrix} 0 & 1 \end{pmatrix}$$

$\square$

3.3. Let A be an invertible matrix, and let $A = W\Sigma V^*$ be its singular value decomposition. Find a singular value decomposition for A^* and A^{-1}.

Answer. Observe that

$$A^* = (W\Sigma V^*)^* = (V^*)^*\Sigma^* W^* = V\Sigma W^*$$

where $\Sigma^* = \Sigma$ since all singular values are real numbers. Also observe that if Σ^{-1} is the matrix equal to Σ except with all diagonal entries inverted (which leaves them as real numbers), then

$$(W\Sigma V^*)(V\Sigma^{-1}W^*) = I \qquad\qquad (V\Sigma^{-1}W^*)(W\Sigma V^*) = I$$

Thus, we have that

$$A^* = V\Sigma W^* \qquad\qquad A^{-1} = V\Sigma^{-1}W^*$$

$\square$

3.5. Find the singular value decomposition of the matrix

$$A = \begin{pmatrix} 2 & 3 \\ 0 & 2 \end{pmatrix}$$

Answer. We have from Problem 6.3.2 that a Schmidt decomposition of A is

$$A = 4 \begin{pmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{pmatrix} \begin{pmatrix} 1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix} + 1 \begin{pmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix} \begin{pmatrix} -2/\sqrt{5} & 1/\sqrt{5} \end{pmatrix}$$

Thus, the singular value decomposition is

$$A = \begin{pmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix} \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ -2/\sqrt{5} & 1/\sqrt{5} \end{pmatrix}$$

$\square$

Use it to find

a) $\max_{\|\mathbf{x}\|\leq 1} \|A\mathbf{x}\|$ and the vectors where the maximum is attained.

Answer. We have that $\max_{\|\mathbf{x}\|\leq 1} \|A\mathbf{x}\| = \|A\| = 4$. We know that the unit vector that maximizes Σ is $\pm\mathbf{e}_1$, so we want to find $\mathbf{x}$ such that $V^*\mathbf{x} = \pm\mathbf{e}_1$. But then $\mathbf{x} = \pm V\mathbf{e}_1$, i.e., $\mathbf{x}$ equals plus or minus the first column of V. Therefore, the vectors where the maximum is attained are

$$\mathbf{x}_1 = \begin{pmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix} \qquad\qquad \mathbf{x}_2 = \begin{pmatrix} -1/\sqrt{5} \\ -2/\sqrt{5} \end{pmatrix}$$

$\square$

b) $\min_{\|\mathbf{x}\|=1} \|A\mathbf{x}\|$ and the vectors where the minimum is attained.

Answer. By a similar argument to before, $\min_{\|\mathbf{x}\|=1} \|A\mathbf{x}\| = 1$ and

$$\mathbf{y}_1 = \begin{pmatrix} -2/\sqrt{5} \\ 1/\sqrt{5} \end{pmatrix} \qquad \mathbf{y}_2 = \begin{pmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{pmatrix}$$

□

c) The image $A(B)$ of the closed unit ball $B = \{\mathbf{x} \in \mathbb{R}^2 : \|\mathbf{x}\| \leq 1\}$ in $\mathbb{R}^2$. Describe $A(B)$ geometrically.

Answer. $A(B)$ will be an ellipse in $\mathbb{R}^2$ centered at the origin with half-axes of length 4 and 1 pointing in the directions $\mathbf{x}_1, \mathbf{x}_2$ and $\mathbf{y}_1, \mathbf{y}_2$, respectively. □

3.6. Show that for a square matrix A, $|\det A| = \det |A|$.

Answer. By Theorem 6.3.5, $A = U|A|$ where U is an isometry. Note that U is unitary in this case as well since U is square (see Proposition 5.6.3). Also note that $\det |A|$ is nonnegative since every eigenvalue of $|A|$ (i.e., the singular values) are nonnegative by definition. Thus,

$$\begin{aligned}
|\det A| &= |\det(U|A|)| \\
&= |\det U| \cdot |\det |A|| && \text{Theorem 3.3.5} \\
&= 1 \cdot |\det |A|| && \text{Proposition 5.6.4} \\
&= \det |A|
\end{aligned}$$

as desired. □

3.7. True or false:

a) The singular values of a matrix are also eigenvalues of the matrix.

Answer. False.
Consider the left matrix in Problem 6.3.2. Since this matrix is upper triangular, it is clear that it's eigenvalue is 2. However, we computed its singular values to be 4 and 1. □

b) The singular values of a matrix A are eigenvalues of A^*A.

Answer. False.
Consider the left matrix in Problem 6.3.2. By the diagonalization of A^*A performed in the answer to that question, the eigenvalues of A^*A are 16 and 1. However, we computed its singular values to be 4 and 1. □

c) If s is a singular value of a matrix A and c is a scalar, then $|c|s$ is a singular value of cA.

Answer. True.
Suppose s is a singular value of A. Then s^2 is an eigenvalue of A^*A, i.e., there exists a nonzero vector $\mathbf{v}$ such that $A^*A\mathbf{v} = s^2\mathbf{v}$. It follows that

$$(cA)^*(cA)\mathbf{v} = c^2 A^*A\mathbf{v} = c^2 s^2 \mathbf{v}$$

so c^2s^2 is an eigenvalue of $(cA)^*(cA)$. Therefore, $\sqrt{c^2s^2} = |c|s$ is a singular value of cA, as desired. □

d) The singular values of any linear operator are nonnegative.

Answer. True.
By definition. □

e) The singular values of a self-adjoint matrix coincide with its eigenvalues.

Answer. False.
Consider the self-adjoint 1×1 matrix

$$A = \begin{pmatrix} -1 \end{pmatrix}$$

The eigenvalue of A is -1, but the singular value is 1. ☐

3.8. Let A be an $m \times n$ matrix. Prove that *nonzero* eigenvalues of the matrices A^*A and AA^* (counting multiplicities) coincide. Can you say when zero eigenvalues of A^*A and zero eigenvalues of AA^* have the same multiplicity?

Answer. Let A be an $m \times n$ matrix with SVD $A = W\Sigma V^*$, and let $\sigma_1, \dots, \sigma_n$ be the singular values of A arranged such that $\sigma_1, \dots, \sigma_r$ are the nonzero singular values. Then

$$\begin{aligned} A^*A &= (W\Sigma V^*)^*(W\Sigma V^*) \\ &= (V^*)^*\Sigma^* W^* W \Sigma V^* \\ &= V\Sigma^*\Sigma V^* \end{aligned} \qquad\qquad \begin{aligned} AA^* &= (W\Sigma V^*)(W\Sigma V^*)^* \\ &= W\Sigma V^* V \Sigma^* W^* \\ &= W\Sigma\Sigma^* W^* \end{aligned}$$

Let's investigate the structure of $\Sigma^*\Sigma$ and $\Sigma\Sigma^*$. By definition, Σ is of the form

$$\begin{array}{r} \\ 1 \\ \vdots \\ r \\ r+1 \\ \vdots \\ m \end{array} \begin{array}{c} \begin{array}{cccccc} 1 & \cdots & r & r+1 & \cdots & n \end{array} \\ \left(\begin{array}{ccc|ccc} \sigma_1 & & & & & \\ & \ddots & & & 0 & \\ & & \sigma_r & & & \\ \hline & & & & & \\ & 0 & & & 0 & \\ & & & & & \end{array}\right) \end{array}$$

It is thus easy to see that

$$\Sigma^*\Sigma = \begin{array}{r} \\ 1 \\ \vdots \\ r \\ r+1 \\ \vdots \\ n \end{array} \begin{array}{c} \begin{array}{cccccc} 1 & \cdots & r & r+1 & \cdots & n \end{array} \\ \left(\begin{array}{ccc|ccc} \sigma_1^2 & & & & & \\ & \ddots & & & 0 & \\ & & \sigma_r^2 & & & \\ \hline & & & & & \\ & 0 & & & 0 & \\ & & & & & \end{array}\right) \end{array} \qquad \Sigma\Sigma^* = \begin{array}{r} \\ 1 \\ \vdots \\ r \\ r+1 \\ \vdots \\ m \end{array} \begin{array}{c} \begin{array}{cccccc} 1 & \cdots & r & r+1 & \cdots & m \end{array} \\ \left(\begin{array}{ccc|ccc} \sigma_1^2 & & & & & \\ & \ddots & & & 0 & \\ & & \sigma_r^2 & & & \\ \hline & & & & & \\ & 0 & & & 0 & \\ & & & & & \end{array}\right) \end{array}$$

i.e., that $\Sigma^*\Sigma$ and $\Sigma\Sigma^*$ are proper diagonal matrices whose entries are the squares of the singular values. This combined with the fact that V, W are unitary means that $V(\Sigma^*\Sigma)V^*$ and $W(\Sigma\Sigma^*)W$ are orthogonal diagonalizations of A^*A and AA^*, respectively. Hence the diagonal entries of $\Sigma^*\Sigma$ are the eigenvalues of A^*A and the diagonal entries of $\Sigma\Sigma^*$ are the the eigenvalues of AA^*. Therefore, from the last equations above, it is clear that the nonzero eigenvalues of A^*A and AA^* always coincide, and the zero eigenvalues of A^*A and AA^* coincide iff $m = n$. ☐

3.9. Let s be the largest singular value of an operator A, and let λ be the eigenvalue of A with the largest absolute value. Show that $|\lambda| \leq s$.

Answer. Let $\mathbf{v}$ be the normal eigenvector corresponding to λ. Then we have that

$$|\lambda| = |\lambda|\|\mathbf{v}\| = \|\lambda\mathbf{v}\| = \|A\mathbf{v}\| \leq \|A\| \cdot \|\mathbf{v}\| = s$$

as desired. ☐

3.11. Show that the operator norm of a matrix A coincides with its Frobenius norm if and only if the matrix has rank one. (Hint: The previous problem might help.)

Answer. Let $\sigma_1, \dots, \sigma_n$ be the singular values of A arranged in descending order.

Suppose first that the $\|A\| = \|A\|_2$. Then

$$\sigma_1^2 = \|A\|^2 = \|A\|_2^2 = \operatorname{tr}(A^*A) = \sum_{k=1}^n \sigma_k^2$$

It follows that $\sigma_2, \dots, \sigma_n$ are all zero. Therefore, since A only has one nonzero singular value, Problem 6.3.1 asserts that A has rank one.

The proof is symmetric in the other direction. ☐

3.12. For the matrix

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

describe the inverse image of the unit ball, i.e., the set of all $\mathbf{x} \in \mathbb{R}^2$ such that $\|A\mathbf{x}\| \leq 1$. Use its singular value decomposition.

Answer. The inverse image of the unit ball under A is equal to the image of the unit ball under A^{-1}. We have that

$$A^{-1} = \frac{1}{4} \begin{pmatrix} 2 & 3 \\ 0 & 2 \end{pmatrix}$$

Thus, by problem 6.3.5, the SVD of A^{-1} is

$$A^{-1} = \begin{pmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1/4 \end{pmatrix} \begin{pmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ -2/\sqrt{5} & 1/\sqrt{5} \end{pmatrix}$$

Thus, the inverse image will be an ellipse in $\mathbb{R}^2$ with half-axes 1 and $\frac{1}{4}$ pointing in the directions $\left(\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} -2 \\ 1 \end{smallmatrix}\right)$, respectively. ☐

4.2. Let A be a normal operator, and let $\lambda_1, \dots, \lambda_n$ be its eigenvalues (counting multiplicities). Show that singular values of A are $|\lambda_1|, \dots, |\lambda_n|$.

Answer. Since A is normal, we have by Theorem 6.2.4 that $A = UDU^*$ where U is unitary and D is diagonal. It follows that

$$A^*A = (UDU^*)^*(UDU^*) = UD^*DU^*$$

Consider λ_j for some $j \in \{1, \dots, n\}$. We know that λ_j is a diagonal entry of D. Thus, $\bar{\lambda_j}\lambda_j = |\lambda_j|^2$ is the corresponding diagonal entry of D^*D. It follows since the singular values of A are the eigenvalues of $|A| = \sqrt{A^*A}$, i.e., the square roots of the eigenvalues of A^*A that $\sigma_j = \sqrt{|\lambda_j|^2} = |\lambda_j|$, as desired. ☐

4.3. Find the singular values, norm, and condition number of the matrix

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

You can do this problem with practically no computations if you use the previous problem and can answer the following questions:

a) What are singular values (eigenvalues) of an orthogonal projection P_E onto some subspace E?

Answer. 1 and 0, with respective multiplicities $\dim E$ and $\dim E^\perp$. Note that the singular values and eigenvalues coincide here because P_E is self-adjoint. ☐

b) What is the matrix of the orthogonal projection onto the subspace spanned by the vector $(1, 1, 1)^T$?

Answer. From Problem 5.3.9a, the matrix of this projection is

$$\frac{1}{3}\begin{pmatrix}1 & 1 & 1\\ 1 & 1 & 1\\ 1 & 1 & 1\end{pmatrix}$$

□

c) How are the eigenvalues of the operators T and $aT + bI$ where $a, b \in \mathbb{F}$ related?

Answer. Suppose λ is an eigenvalue of T. Then there exists a nonzero vector $\mathbf{v}$ such that $T\mathbf{v} = \lambda\mathbf{v}$. It follows that

$$\begin{aligned}(aT + bI)\mathbf{v} &= aT\mathbf{v} + b\mathbf{v}\\ &= a\lambda\mathbf{v} + b\mathbf{v}\\ &= (a\lambda + b)\mathbf{v}\end{aligned}$$

i.e., that $a\lambda + b$ is an eigenvalue of $aT + bI$. □

Of course you can also just honestly do the computations.

Answer. Let P_E denote the matrix provided as an answer to question (b) above. Then $A = 3P_E + 1I$. Therefore, since question (a) provides the eigenvalues to P_E as 1 and 0 (with multiplicities 2 and 1, respectively), question (c) posits that the eigenvalues of A are $3(1) + 1 = 4$ and $3(0) + 1 = 1$ (with multiplicities 2 and 1, respectively), and that these values are in fact the singular values.

It follows that $\|A\| = 4$ and the condition number is $\|A\| \cdot \|A^{-1}\| = 4/1 = 4$. □

6.1. Let R_α be the rotation through α, so its matrix in the standard basis is

$$\begin{pmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{pmatrix}$$

Find the matrix of R_α in the basis $\mathbf{v}_1, \mathbf{v}_2$ where $\mathbf{v}_1 = \mathbf{e}_2, \mathbf{v}_2 = \mathbf{e}_1$.

Answer. We define

$$[I]_{\mathcal{E}\mathcal{V}} = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}$$

It follows that

$$[I]_{\mathcal{V}\mathcal{E}} = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}$$

Therefore,

$$\begin{aligned}[R_\alpha]_{\mathcal{V}\mathcal{V}} &= [I]_{\mathcal{V}\mathcal{E}}[R_\alpha]_{\mathcal{E}\mathcal{E}}[I]_{\mathcal{E}\mathcal{V}}\\ &= \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{pmatrix}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\\ &= \begin{pmatrix}\cos\alpha & \sin\alpha\\ -\sin\alpha & \cos\alpha\end{pmatrix}\end{aligned}$$

□

6.2. Let R_α be the rotation matrix

$$R_\alpha = \begin{pmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{pmatrix}$$

Show that the 2×2 identity matrix I_2 can be continuously transformed through invertible matrices into R_α.

Answer. Let $V(t)$ be defined by

$$V(t) = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$$

Choose $a = 0$ and $b = \alpha$. Then $V(t)$ is continuous because each component is continuous, the inverse of $V(t)$ is $V(-t)$, and clearly $V(a) = V(0) = I$ and $V(b) = V(\alpha) = R_\alpha$. $\square$

6.3. Let U be an $n \times n$ orthogonal matrix with $\det U > 0$. Show that the $n \times n$ identity matrix I_n can be continuously transformed through invertible matrices into U. (Hint: Use the previous problem and the representation of a rotation in $\mathbb{R}^n$ as a product of planar rotations [see Section 5].)

Answer. Since U is an orthogonal matrix with $\det U = 1 > 0$, Theorem 6.5.1 asserts that there exists a basis $\mathbf{v}_1, \dots, \mathbf{v}_n$ such that the matrix of U in this basis has the block diagonal form

$$V(t) = \begin{pmatrix} R_{\varphi_1} & & & 0 \\ & \ddots & & \\ & & R_{\varphi_k} & \\ 0 & & & I_{n-2k} \end{pmatrix}$$

Thus, let $V(t)$ be defined by

$$V(t) = \begin{pmatrix} R_{\varphi_1 t} & & & 0 \\ & \ddots & & \\ & & R_{\varphi_k t} & \\ 0 & & & I_{n-2k} \end{pmatrix}$$

Choose $a = 0$ and $b = 1$. It will follow from Problem 6.6.2 that V is a continuous transformation satisfying all the necessary properties. $\square$

Chapter 7

1.1. Find the matrix of the bilinear form L on $\mathbb{R}^3$ defined by

$$L(\mathbf{x}, \mathbf{y}) = x_1y_1 + 2x_1y_2 + 14x_1y_3 - 5x_2y_1 + 2x_2y_2 - 3x_2y_3 + 8x_3y_1 + 19x_3y_2 - 2x_3y_3$$

Answer. We have that

$$\begin{aligned}
&= x_1y_1 + 2x_1y_2 + 14x_1y_3 - 5x_2y_1 + 2x_2y_2 - 3x_2y_3 + 8x_3y_1 + 19x_3y_2 - 2x_3y_3 \\
&= L(\mathbf{x}, \mathbf{y}) \\
&= \mathbf{y}^T A\mathbf{x} \\
&= \begin{pmatrix} y_1 & y_2 & y_3 \end{pmatrix} \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\
&= \begin{pmatrix} y_1 & y_2 & y_3 \end{pmatrix} \begin{pmatrix} a_{1,1}x_1 + a_{1,2}x_2 + a_{1,3}x_3 \\ a_{2,1}x_1 + a_{2,2}x_2 + a_{2,3}x_3 \\ a_{3,1}x_1 + a_{3,2}x_2 + a_{3,3}x_3 \end{pmatrix} \\
&= y_1a_{1,1}x_1 + y_1a_{1,2}x_2 + y_1a_{1,3}x_3 + y_2a_{2,1}x_1 + y_2a_{2,2}x_2 + y_2a_{2,3}x_3 + y_3a_{3,1}x_1 + y_3a_{3,2}x_2 + y_3a_{3,3}x_3 \\
&= a_{1,1}x_1y_1 + a_{2,1}x_1y_2 + a_{3,1}x_1y_3 + a_{1,2}x_2y_1 + a_{2,2}x_2y_2 + a_{3,2}x_2y_3 + a_{1,3}x_3y_1 + a_{2,3}x_3y_2 + a_{3,3}x_3y_3
\end{aligned}$$

It follows from comparing terms that

$$A = \begin{pmatrix} 1 & -5 & 8 \\ 2 & 2 & 19 \\ 14 & -3 & -2 \end{pmatrix}$$

$\square$

1.2. Define the bilinear form L on $\mathbb{R}^2$ by

$$L(\mathbf{x},\mathbf{y}) = \det[\mathbf{x},\mathbf{y}]$$

i.e., to compute $L(\mathbf{x},\mathbf{y})$, we form a 2×2 matrix with columns $\mathbf{x},\mathbf{y}$ and compute its determinant. Find the matrix of L.

Answer. We have that

$$\begin{aligned}
&= a_{1,1}x_1y_1 + a_{2,1}x_1y_2 + a_{1,2}x_2y_1 + a_{2,2}x_2y_2 \\
&= \mathbf{y}^T A\mathbf{x} \\
&= L(\mathbf{x},\mathbf{y}) \\
&= \det\begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix} \\
&= x_1y_2 - y_1x_2 \\
&= 0x_1y_1 + 1x_1y_2 - 1x_2y_1 + 0x_2y_2
\end{aligned}$$

It follows from comparing terms that

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

$\square$

1.3. Find the matrix of the quadratic form Q on $\mathbb{R}^3$ defined by

$$Q[\mathbf{x}] = x_1^2 + 2x_1x_2 - 3x_1x_3 - 9x_2^2 + 6x_2x_3 + 13x_3^2$$

Answer. We have that

$$\begin{aligned}
&= x_1^2 + 2x_1x_2 - 3x_1x_3 - 9x_2^2 + 6x_2x_3 + 13x_3^2 \\
&= Q[\mathbf{x}] \\
&= (A\mathbf{x},\mathbf{x}) \\
&= \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{1,2} & a_{2,2} & a_{2,3} \\ a_{1,3} & a_{2,3} & a_{3,3} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\
&= \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} a_{1,1}x_1 + a_{1,2}x_2 + a_{1,3}x_3 \\ a_{1,2}x_1 + a_{2,2}x_2 + a_{2,3}x_3 \\ a_{1,3}x_1 + a_{2,3}x_2 + a_{3,3}x_3 \end{pmatrix} \\
&= x_1a_{1,1}x_1 + x_1a_{1,2}x_2 + x_1a_{1,3}x_3 + x_2a_{1,2}x_1 + x_2a_{2,2}x_2 + x_2a_{2,3}x_3 + x_3a_{1,3}x_1 + x_3a_{2,3}x_2 + x_3a_{3,3}x_3 \\
&= a_{1,1}x_1^2 + 2a_{1,2}x_1x_2 + 2a_{1,3}x_1x_3 + a_{2,2}x_2^2 + 2a_{2,3}x_2x_3 + a_{3,3}x_3^2
\end{aligned}$$

It follows from comparing terms that

$$A = \begin{pmatrix} 1 & 1 & -3/2 \\ 1 & -9 & 3 \\ -3/2 & 3 & 13 \end{pmatrix}$$

$\square$

2.1. Diagonalize the quadratic form with the matrix

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 1 \end{pmatrix}$$

Use two methods: completion of squares and row operations. Which one do you like better? Can you say if the matrix A is positive definite or not?

Answer. Completion of squares: If A has the above form, then

$$\begin{aligned}
Q[\mathbf{x}] &= (A\mathbf{x}, \mathbf{x}) \\
&= a_{1,1}x_1^2 + 2a_{1,2}x_1x_2 + 2a_{1,3}x_1x_3 + a_{2,2}x_2^2 + 2a_{2,3}x_2x_3 + a_{3,3}x_3^2 \\
&= x_1^2 + 4x_1x_2 + 2x_1x_3 + 3x_2^2 + 4x_2x_3 + x_3^2 \\
&= (x_1 + 2x_2 + x_3)^2 - x_2^2 \\
&= y_1^2 - y_2^2
\end{aligned}$$

where $y_1 = x_1 + 2x_2 + x_3$, $y_2 = x_2$, and $y_3 = 0$. It follows that

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad S^* = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

the latter equation coming from the fact that $\mathbf{y} = S^*\mathbf{x}$.

Row operations: We can row reduce

$$(A|I) = \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 2 & 3 & 2 & 0 & 1 & 0 \\ 1 & 2 & 1 & 0 & 0 & 1 \end{array}\right)$$

to

$$(D|S^*) = \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 2 & 1 \\ 0 & -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

getting the same result as before.

Right now, I believe I prefer completion of squares. The matrix is not positive definite since it has an eigenvalue (diagonal entry) less than zero. □

5 Definiteness, Dual Spaces, and Advanced Spectral Theory

From Treil (2017).

Chapter 7

11/1: **4.1.** Using Silvester's Criterion of Positivity, check if the matrices

$$A = \begin{pmatrix} 4 & 2 & 1 \\ 2 & 3 & -1 \\ 1 & -1 & 2 \end{pmatrix} \qquad B = \begin{pmatrix} 3 & -1 & 2 \\ -1 & 4 & -2 \\ 2 & -2 & 2 \end{pmatrix}$$

are positive definite or not. Are the matrices $-A$, A^3, A^{-1}, $A + B^{-1}$, $A + B$, and $A - B$ positive definite?

Answer. $\underline{A}$: We have that

$$\det A_1 = \det \begin{pmatrix} 4 \end{pmatrix} = 4 \qquad \det A_2 = \det \begin{pmatrix} 4 & 2 \\ 2 & 3 \end{pmatrix} = 8 \qquad \det A_3 = \det \begin{pmatrix} 4 & 2 & 1 \\ 2 & 3 & -1 \\ 1 & -1 & 2 \end{pmatrix} = 5$$

Thus, since $A = A^*$ and $\det A_k > 0$ for $k = 1, 2, 3$, Silvester's Criterion of Positivity implies that A is positive definite.

$\underline{B}$: We have that

$$\det B_1 = \det \begin{pmatrix} 3 \end{pmatrix} = 3 \qquad \det B_2 = \det \begin{pmatrix} 3 & -1 \\ -1 & 4 \end{pmatrix} = 11 \qquad \det B_3 = \det \begin{pmatrix} 3 & -1 & 2 \\ -1 & 4 & -2 \\ 2 & -2 & 2 \end{pmatrix} = 2$$

Thus, since $B = B^*$ and $\det B_k > 0$ for $k = 1, 2, 3$, Silvester's Criterion of Positivity implies that B is positive definite.

$\underline{-A}$: We have that

$$\det(-A)_1 = \det \begin{pmatrix} -4 \end{pmatrix} = -4 \not> 0$$

Thus, Silvester's Criterion of Positivity implies that B is not positive definite.

$\underline{A^3}$: Since $A = A^*$, Theorem 6.2.2 implies that $A = UDU^*$ where U is unitary and $D = \operatorname{diag}\{\lambda_1, \lambda_2, \lambda_3\}$ with each λ_k real. Moreover, since A is positive definite, Theorem 7.4.1 implies that each $\lambda_k > 0$. Thus, since $A^3 = UD^3U^*$, A^3 is Hermitian, $D^3 = \operatorname{diag}\{\lambda_1^3, \lambda_2^3, \lambda_3^3\}$ where each λ_k^3 is an eigenvalue of A^3, and naturally each $\lambda_k^3 > 0$, Theorem 7.4.1 implies that A^3 is positive definite.

$\underline{A^{-1}}$: By a symmetric argument to the one used for A^3, we have that A^{-1} is positive definite.

$\underline{A + B^{-1}}$: Since A is positive definite, by definition, $(A\mathbf{x}, \mathbf{x}) > 0$ for all $\mathbf{x} \neq \mathbf{0}$. By a symmetric argument to the one used for A^{-1}, B^{-1} is positive definite. Thus, similarly, $(B^{-1}\mathbf{x}, \mathbf{x}) > 0$ for all $\mathbf{x} \neq \mathbf{0}$. It follows by combining the previous results that if $\mathbf{x} \neq \mathbf{0}$, then

$$0 < (A\mathbf{x}, \mathbf{x}) < (A\mathbf{x}, \mathbf{x}) + (B^{-1}\mathbf{x}, \mathbf{x}) = ((A + B^{-1})\mathbf{x}, \mathbf{x})$$

so $A + B^{-1}$ is positive definite.

$\underline{A + B}$: By a symmetric argument to the one used for $A + B^{-1}$, we have that $A + B$ is positive definite.

$\underline{A - B}$: We have that

$$\det(A - B)_2 = \det \begin{pmatrix} 1 & 3 \\ 3 & -1 \end{pmatrix} = -10 \not> 0$$

Thus, Silvester's Criterion of Positivity implies that $A - B$ is not positive definite. ☐

4.2. True or false:

a) If A is positive definite, then A^5 is positive definite.

Answer. True.
If A is positive definite, then $A = A^*$. It follows that $A = UDU^*$. Additionally, Theorem 7.4.1 implies that $\lambda_k > 0$ for all λ_k along the diagonal of D. Thus, $A^5 = UD^5U^*$ where D^5 has all positive diagonal entries because D has all positive diagonal entries. Thus, by Theorem 7.4.1 again, A^5 is positive definite. □

b) If A is negative definite, then A^8 is negative definite.

Answer. False.
If A is negative definite, then as before, $A = UDU^*$ and $A^8 = UD^8U^*$. But if every entry along the diagonal of D is negative (Theorem 7.4.1), then every diagonal along $D^8 = (D^2)^4$ will be positive, so A^8 is not negative definite (it is, in fact, positive definite). □

c) If A is negative definite, then A^{12} is positive definite.

Answer. True.
See the explanation to part (b). □

d) If A is positive definite and B is negative semidefinite, then $A - B$ is positive definite.

Answer. True.
If A is positive definite, then $(A\mathbf{x}, \mathbf{x}) > 0$ for all $\mathbf{x} \neq 0$. Similarly, $(B\mathbf{x}, \mathbf{x}) \leq 0$ for all $\mathbf{x}$. To prove that $A - B$ is positive definite, it will suffice to show that $((A - B)\mathbf{x}, \mathbf{x}) > 0$ for all $\mathbf{x} \neq 0$. Let $\mathbf{x} \neq 0$ be arbitrary. Then

$$0 < (A\mathbf{x}, \mathbf{x}) \leq (A\mathbf{x}, \mathbf{x}) - (B\mathbf{x}, \mathbf{x}) = (A\mathbf{x} - B\mathbf{x}, \mathbf{x}) = ((A - B)\mathbf{x}, \mathbf{x})$$

as desired. □

e) If A is indefinite, and B is positive definite, then $A + B$ is indefinite.

Answer. False.
Let

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

By Theorem 7.4.1, A is indefinite and B is positive definite. However,

$$A + B = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$$

which is positive semidefinite by Theorem 7.4.1. □

4.3. Let A be a 2×2 Hermitian matrix such that $a_{1,1} > 0$, $\det A \geq 0$. Prove that A is positive semidefinite.

Answer. We have by the given constraints that A is of the form

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} \\ \bar{a}_{1,2} & a_{2,2} \end{pmatrix}$$

Additionally, we have that

$$\begin{aligned} 0 \leq \det A &= a_{1,1}a_{2,2} - a_{1,2}\bar{a}_{1,2} = a_{1,1}a_{2,2} - |a_{1,2}|^2 \\ |a_{1,2}|^2 &\leq a_{1,1}a_{2,2} \end{aligned}$$

from which it follows since $|a_{1,2}|^2 \geq 0$ that

$$0 \leq |a_{1,2}|^2 \leq a_{1,1}a_{2,2}$$

This combined with the fact that $a_{1,1} > 0$ implies that $a_{2,2} \geq 0$. Thus,

$$\operatorname{tr} A = a_{1,1} + a_{2,2} \geq a_{1,1} + 0 > 0$$

Now let λ_1, λ_2 be the eigenvalues of A. It follows from the above since $\operatorname{tr} A = \lambda_1 + \lambda_2$ that WLOG we may let $\lambda_1 > 0$. It follows that

$$\begin{aligned} 0 &\leq \det A = \lambda_1\lambda_2 \\ 0 &\leq \lambda_2 \end{aligned}$$

Therefore, having shown that each $\lambda_k \geq 0$, Theorem 7.4.1 implies that A is positive semidefinite, as desired. □

4.4. Find a real symmetric $n \times n$ matrix such that $a_{1,1} > 0$ and $\det A_k \geq 0$ for all $k = 2, \dots, n$, but the matrix A is not positive semidefinite. Try to find an example for the minimal possible n[1].

Answer. Let

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Then $a_{1,1} = 1 > 0$, $\det A_2 = 0 \geq 0$, and $\det A_3 = 0 \geq 0$. However, we have that its eigenvalues are $\lambda = -1, 0, 2$, so A is actually indefinite. Also, we know that this is the answer for the minimal possible n since Problem 7.4.3 proves that the conditions actually *do* imply A is positive semidefinite for $n = 2$. □

4.5. Let A be an $n \times n$ Hermitian matrix such that $\det A_k > 0$ for all $k = 1, \dots, n-1$ and $\det A \geq 0$. Prove that A is positive semidefinite.

Answer. Let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of A, and let $\mu_1, \dots, \mu_{n-1}$ be the eigenvalues of A_{n-1}, both sets taken in decreasing order. By Silvester's Criterion of Positivity, the hypothesis that $\det A_k > 0$ for each $k = 1, \dots, n-1$ implies that A_{n-1} is positive definite. Thus, by Theorem 7.4.1, each $\mu_k > 0$. It follows by Corollary 7.4.4 that

$$\lambda_k \geq \mu_{n-1} > 0$$

for each $k = 1, \dots, n-1$. Thus,

$$\begin{aligned} 0 &\leq \det A = \lambda_1 \cdots \lambda_{n-1}\lambda_n \\ 0 &\leq \lambda_n \end{aligned}$$

Therefore, since each $\lambda_k \geq 0$, Theorem 7.4.1 implies that A is positive semidefinite, as desired. □

4.6. Find a real symmetric 3×3 matrix A such that $a_{1,1} > 0$, $\det A_k \geq 0$ for $k = 2, 3$, but the matrix A is not positive semidefinite.

Answer. Using the same matrix from Problem 7.4.4, we have

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

□

[1] The statement of this problem has been modified as per Chloé's instructions in the 10/28 problem session.

Chapter 8

1.1. Let $\mathbf{v}_1, \dots, \mathbf{v}_r$ be a system of vectors in X such that there exists a system $\mathbf{v}_1', \dots, \mathbf{v}_r'$ of linear functionals such that

$$\mathbf{v}_k'(\mathbf{v}_j) = \delta_{jk}$$

a) Show that the system $\mathbf{v}_1, \dots, \mathbf{v}_r$ is linearly independent.

Answer. To prove that $\mathbf{v}_1, \dots, \mathbf{v}_r$ is linearly independent, it will suffice to show that if $\alpha_1, \dots, \alpha_r \in \mathbb{F}$ make $\alpha_1\mathbf{v}_1 + \cdots + \alpha_r\mathbf{v}_r = 0$, then $\alpha_1 = \cdots = \alpha_r = 0$. Suppose that $\alpha_1, \dots, \alpha_r \in \mathbb{F}$ make

$$\alpha_1\mathbf{v}_1 + \cdots + \alpha_r\mathbf{v}_r = 0$$

It follows by linearity and the definition of the dual basis that

$$\begin{aligned}
0 &= \mathbf{v}_k'(\mathbf{0}) \\
&= \mathbf{v}_k'(\alpha_1\mathbf{v}_1 + \cdots + \alpha_r\mathbf{v}_r) \\
&= \alpha_1\mathbf{v}_k'(\mathbf{v}_1) + \cdots + \alpha_r\mathbf{v}_k'(\mathbf{v}_r) \\
&= \alpha_1 \cdot 0 + \cdots + \alpha_{k-1} \cdot 0 + \alpha_k \cdot 1 + \alpha_{k+1} \cdot 0 + \cdots + \alpha_r \cdot 0 \\
&= \alpha_k
\end{aligned}$$

for each $k = 1, \dots, r$, as desired. $\square$

b) Show that if the system $\mathbf{v}_1, \dots, \mathbf{v}_r$ is not generating, then the "biorthogonal" system $\mathbf{v}_1', \dots, \mathbf{v}_r'$ is not unique. (Hint: Probably the easiest way to prove that is to complete the system $\mathbf{v}_1, \dots, \mathbf{v}_r$ to a basis [see Proposition 2.5.4].)

Answer. By Proposition 2.5.4, we can complete the linearly independent list $\mathbf{v}_1, \dots, \mathbf{v}_r$ to a basis $\mathbf{v}_1, \dots, \mathbf{v}_n$ where $n > r$ since $\mathbf{v}_1, \dots, \mathbf{v}_r$ is not generating by hypothesis. Consider $\mathbf{v}_1', \dots, \mathbf{v}_r'$. These linear functionals' behavior on $\mathbf{v}_1, \dots, \mathbf{v}_r$ is completely defined by the given condition; however, since they act on all of X and not just $\operatorname{span}\{\mathbf{v}_1, \dots, \mathbf{v}_r\} \subsetneq X$, we can define an arbitrary linear behavior for each $\mathbf{v}_k'$ on $\operatorname{span}\{\mathbf{v}_{r+1}, \dots, \mathbf{v}_n\}$. Clearly, more than one such behavior exists (take, for example, being the zero map on that span and being the identity map on that span), so $\mathbf{v}_1', \dots, \mathbf{v}_n'$ is not unique. $\square$

3.1. Prove that if for linear transformations $T, T_1 : X \to Y$

$$\langle T\mathbf{x}, \mathbf{y}' \rangle = \langle T_1\mathbf{x}, \mathbf{y}' \rangle$$

for all $\mathbf{x} \in X$ and for all $\mathbf{y}' \in Y'$, then $T = T_1$. (Hint: Probably one of the easiest ways of proving this is to use Lemma 8.1.3.)

Answer. Let $\mathbf{x} \in X$ be arbitrary. If $\langle T\mathbf{x}, \mathbf{y}' \rangle = \langle T_1\mathbf{x}, \mathbf{y}' \rangle$ for all $\mathbf{y}' \in Y'$, then $\mathbf{y}'(T\mathbf{x}) = \mathbf{y}'(T_1\mathbf{x})$ for all $\mathbf{x} \in X$ and for all $\mathbf{y}' \in Y'$. Thus, since every linear functional in the dual space maps the vectors $T\mathbf{x}$ and $T_1\mathbf{x}$ the same way, Lemma 8.1.3 implies that $T\mathbf{x} = T_1\mathbf{x}$. But since we let $\mathbf{x}$ be arbitrary, $T\mathbf{x} = T_1\mathbf{x}$ for all $\mathbf{x} \in X$, i.e., $T = T_1$. $\square$

3.2. Combine the Riesz Representation Theorem (Theorem 8.2.1) with the reasoning in Section 3.1.3 above to present a coordinate-free definition of the Hermitian adjoint of an operator in an inner product space.

Answer. Let $A \in \mathcal{L}(V, W)$. We seek to define A^* as the unique element of $\mathcal{L}(W, V)$ satisfying

$$(A\mathbf{x}, \mathbf{y}) = (\mathbf{x}, A^*\mathbf{y})$$

for all $\mathbf{x} \in V$ and $\mathbf{y} \in W$. Let's begin.

Let $\mathbf{y}$ be an arbitrary element of W. We can think of $\mathbf{y}^*$ as a $1 \times \dim W$ matrix, or indeed a linear transformation $\mathbf{y}^* : W \to \mathbb{F}$. This combined with the fact that $A : V \to W$ implies that $\mathbf{y}^* A : V \to \mathbb{F}$ is a well-defined linear functional. It follows by the Riesz Representation Theorem that there exists a unique $\mathbf{z} \in V$ such that $(\mathbf{y}^* A)(\mathbf{x}) = (\mathbf{x}, \mathbf{z})$ for all $\mathbf{x} \in V$. Define $A^*\mathbf{y} := \mathbf{z}$.

Since $\mathbf{z}$ is unique by the Riesz Representation Theorem, A^* is a well-defined function for this $\mathbf{y}$. Moreover, since we let $\mathbf{y} \in W$ be arbitrary, we can define $A^*\mathbf{y}$ in the same way for *any* $\mathbf{y} \in W$. Thus, $A^* : W \to Z$ (as defined) is a well-defined function on W.

We now seek to prove that A^* is linear. Let $\mathbf{y}_1, \mathbf{y}_2 \in W$ and $\alpha_1, \alpha_2 \in \mathbb{F}$. We know that $A^*\mathbf{y}_1$ is the unique vector $\mathbf{z}_1 \in V$ such that $(\mathbf{y}_1^* A)(\mathbf{x}) = (\mathbf{x}, \mathbf{z}_1)$ for all $\mathbf{x} \in V$. We also know that $A^*\mathbf{y}_2$ is the unique vector $\mathbf{z}_2 \in V$ such that $(\mathbf{y}_2^* A)(\mathbf{x}) = (\mathbf{x}, \mathbf{z}_2)$ for all $\mathbf{x} \in V$. Lastly, we know that $A^*(\alpha_1\mathbf{y}_1 + \alpha_2\mathbf{y}_2)$ is the unique vector $\mathbf{z} \in V$ such that $[(\alpha_1\mathbf{y}_1 + \alpha_2\mathbf{y}_2)^* A](\mathbf{x}) = (\mathbf{x}, \mathbf{z})$ for all $\mathbf{x} \in V$. It follows that

$$\begin{aligned}
(\mathbf{x}, A^*(\alpha_1\mathbf{y}_1 + \alpha_2\mathbf{y}_2)) &= (\mathbf{x}, \mathbf{z}) \\
&= [(\alpha_1\mathbf{y}_1 + \alpha_2\mathbf{y}_2)^* A](\mathbf{x}) \\
&= \bar{\alpha}_1(\mathbf{y}_1^* A)(\mathbf{x}) + \bar{\alpha}_2(\mathbf{y}_2^* A)(\mathbf{x}) \\
&= \bar{\alpha}_1(\mathbf{x}, \mathbf{z}_1) + \bar{\alpha}_2(\mathbf{x}, \mathbf{z}_2) \\
&= (\mathbf{x}, \alpha_1\mathbf{z}_1 + \alpha_2\mathbf{z}_2) \\
&= (\mathbf{x}, \alpha_1 A^*\mathbf{y}_1 + \alpha_2 A^*\mathbf{y}_2)
\end{aligned}$$

for all $\mathbf{x} \in V$. Thus, by Lemma 8.1.3

$$A^*(\alpha_1\mathbf{y}_1 + \alpha_2\mathbf{y}_2) = \alpha_1 A^*\mathbf{y}_1 + \alpha_2 A^*\mathbf{y}_2$$

as desired.

We now show that A^* satisfies the desired identity: If $\mathbf{x} \in V$ and $\mathbf{y} \in W$, then we have by the definition of A^* that

$$(\mathbf{x}, A^*\mathbf{y}) = (\mathbf{y}^* A)(\mathbf{x}) = \mathbf{y}^*(A\mathbf{x}) = (A\mathbf{x}, \mathbf{y})$$

as desired.

Lastly, we prove that A^* is the unique linear map satisfying the above identity. Suppose $A^*, \tilde{A}^*$ are linear maps such that

$$(A\mathbf{x}, \mathbf{y}) = (\mathbf{x}, A^*\mathbf{y}) \qquad\qquad (A\mathbf{x}, \mathbf{y}) = (\mathbf{x}, \tilde{A}^*\mathbf{y})$$

for all $\mathbf{x} \in V$ and $\mathbf{y} \in W$. Let $\mathbf{y} \in W$ be arbitrary. Then

$$(\mathbf{x}, A^*\mathbf{y}) = (A\mathbf{x}, \mathbf{y}) = (\mathbf{x}, \tilde{A}^*\mathbf{y})$$

for all $\mathbf{x} \in V$. It follows by Lemma 8.1.3 that $A^*\mathbf{y} = \tilde{A}^*\mathbf{y}$. Furthermore, since we let $\mathbf{y}$ be arbitrary, we know that $A^*\mathbf{y} = \tilde{A}^*\mathbf{y}$ for *every* $\mathbf{y} \in W$. Therefore, $A^* = \tilde{A}^*$, so A^* is unique, as desired. ☐

3.3. Let $\mathbf{v}_1, \dots, \mathbf{v}_n$ be a basis in X and let $\mathbf{v}_1', \dots, \mathbf{v}_n'$ be its dual basis. Let $E = \operatorname{span}\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ for $r < n$. Prove that $E^\perp = \operatorname{span}\{\mathbf{v}_{r+1}', \dots, \mathbf{v}_n'\}$. (This problem gives a way to prove Proposition 8.3.6.)

Answer. Suppose first that $\mathbf{v}' \in E^\perp$. Then by the definition of the annihilator, $\mathbf{v}' \in X'$ and $\langle \mathbf{x}, \mathbf{v}' \rangle = 0$ for all $\mathbf{x} \in E$. It follows from the first condition that

$$\mathbf{v}' = \alpha_1\mathbf{v}_1' + \cdots + \alpha_n\mathbf{v}_n'$$

for some $\alpha_1, \dots, \alpha_n \in \mathbb{F}$. It follows from the second condition that

$$\begin{aligned}
0 &= \langle \mathbf{v}_k, \mathbf{v}' \rangle \\
&= \alpha_1\mathbf{v}_1'(\mathbf{v}_k) + \cdots + \alpha_n\mathbf{v}_n'(\mathbf{v}_k) \\
&= \alpha_k
\end{aligned}$$

for each $k = 1, \dots, r$. Therefore,

$$\mathbf{v}' = \alpha_{r+1}\mathbf{v}'_{r+1} + \cdots + \alpha_n \mathbf{v}'_n$$

so $\mathbf{v} \in \operatorname{span}\{\mathbf{v}'_{r+1}, \dots, \mathbf{v}'_n\}$, as desired.

Now suppose that $\mathbf{v}' \in \operatorname{span}\{\mathbf{v}'_{r+1}, \dots, \mathbf{v}'_n\}$. In particular, let $\mathbf{v}' = \alpha_{r+1}\mathbf{v}'_{r+1} + \cdots + \alpha_n \mathbf{v}'_n$ for some $\alpha_{r+1}, \dots, \alpha_n \in \mathbb{F}$. To prove that $\mathbf{v}' \in E^\perp$, it will suffice to show that $\langle \mathbf{x}, \mathbf{v}' \rangle = 0$ for all $\mathbf{x} \in E$. Let $\mathbf{x}$ be an arbitrary element of E. Then by the definition of E, $\mathbf{x} = \beta_1\mathbf{v}_1 + \cdots + \beta_r\mathbf{v}_r$. It follows by the definition of $\mathbf{v}'$ and the dual basis that

$$\begin{aligned}
\langle \mathbf{x}, \mathbf{v}' \rangle &= \alpha_{r+1}\mathbf{v}'_{r+1}(\beta_1\mathbf{v}_1 + \cdots + \beta_r\mathbf{v}_r) + \cdots + \alpha_n\mathbf{v}'_n(\beta_1\mathbf{v}_1 + \cdots + \beta_r\mathbf{v}_r) \\
&= \alpha_{r+1} \cdot 0 + \cdots + \alpha_n \cdot 0 \\
&= 0
\end{aligned}$$

as desired. $\square$

Chapter 9

1.1. (Cayley-Hamilton Theorem for diagonalizable matrices). As discussed in Section 9.1, the Cayley-Hamilton theorem states that if A is a square matrix and

$$p(\lambda) = \det(A - \lambda I) = \sum_{k=0}^{n} c_k \lambda^k$$

is its characteristic polynomial, then $p(A) = \sum_{k=0}^{n} c_k A^k = \mathbf{0}$ (assuming that by definition, $A^0 = I$). Prove this theorem for the special case when A is similar to a diagonal matrix, i.e., $A = SDS^{-1}$. (Hint: If $D = \operatorname{diag}\{\lambda_1, \dots, \lambda_n\}$ and p is any polynomial, can you compute $p(D)$? What about $p(A)$?)

Answer. Suppose $A = SDS^{-1}$, and let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of A. Since $D = \operatorname{diag}\{\lambda_1, \dots, \lambda_n\}$, we have by the properties of diagonal matrix exponentiation, scalar multiplication, and addition, and Exercise 4.1.10 that

$$\begin{aligned}
p(D) &= \sum_{k=0}^{n} c_k D^k \\
&= \sum_{k=0}^{n} c_k \operatorname{diag}\{\lambda_1^k, \dots, \lambda_n^k\} \\
&= \sum_{k=0}^{n} \operatorname{diag}\{c_k\lambda_1^k, \dots, c_k\lambda_n^k\} \\
&= \operatorname{diag}\left\{\sum_{k=0}^{n} c_k\lambda_1^k, \dots, \sum_{k=0}^{n} c_k\lambda_n^k\right\} \\
&= \operatorname{diag}\{p(\lambda_1), \dots, p(\lambda_n)\} \\
&= \operatorname{diag}\{0, \dots, 0\} \\
&= 0
\end{aligned}$$

It follows that

$$\begin{aligned}
p(A) &= p(SDS^{-1}) \\
&= \sum_{k=0}^{n} c_k (SDS^{-1})^k \\
&= \sum_{k=0}^{n} c_k SD^k S^{-1} \\
&= S\left[\sum_{k=0}^{n} c_k D^k\right] S^{-1} \\
&= S[p(D)]S^{-1} \\
&= S0S^{-1} \\
&= 0
\end{aligned}$$

as desired. $\square$

2.1. An operator A is called **nilpotent** if $A^k = \mathbf{0}$ for some k. Prove that if A is nilpotent, then $\sigma(A) = \{0\}$ (i.e., that 0 is the only eigenvalue of A). Can you do it without using the spectral mapping theorem?

Answer. Suppose for the sake of contradiction that $\lambda \neq 0$ for some eigenvalue λ of A. Then if $\mathbf{v}$ is a nonzero eigenvector corresponding to λ, $A\mathbf{v} = \lambda\mathbf{v}$ so $A^k\mathbf{v} = \lambda^k\mathbf{v}$. But since $\lambda^k\mathbf{v} \neq \mathbf{0}$, $A^k \neq 0$, a contradiction. $\square$

6 Basic Topology

From Rudin (1976).

Chapter 2

11/8: **1.** Prove that the empty set is a subset of every set.

Proof. Let A be a set. Then $x \in A$ for all $x \in \emptyset$ is vacuously true. Thus, $\emptyset \subset A$. □

2. A complex number z is said to be **algebraic** if there are integers $a_0, \dots, a_n$, not all zero, such that

$$a_0 z^n + a_1 z^{n-1} + \cdots + a_{n-1} z + a_n = 0$$

Prove that the set of all algebraic numbers is countable. (Hint: For every positive integer N, there are only finitely many equations with $n + |a_0| + |a_1| + \cdots + |a_n| = N$.)

Proof. Define a family of sets $\{A_N\}$ such that each A_N is the set of all complex zeroes of all polynomials $\sum_{k=0}^{n} a_k z^{n-k}$ with integer coefficients $a_0, \dots, a_n$, not all zero, satisfying the equation $n + |a_0| + \cdots + |a_n| = N$. Symbolically, let each

$$A_N = \{z \in \mathbb{C} \mid \textstyle\sum_{k=0}^{n} a_k z^{n-k} = 0,\ a_0, \dots, a_n \in \mathbb{Z},\ \exists\ a_i : a_i \neq 0,\ n + |a_0| + \cdots + |a_n| = N\}$$

Since there are only finitely many equations with $n + |a_0| + \cdots + |a_n| = N$ for each N by the hint, there are only finitely many corresponding polynomials $\sum_{k=0}^{n} a_k z^{n-k}$ for each N. By the fundamental theorem of arithmetic, every polynomial p has at most $\deg p$ distinct solutions. Thus, since each A_N is the union of finitely many finite sets, each A_N is finite.

Consider the set $A = \bigcup_{N=1}^{\infty} A_N$. Since every algebraic number is a zero of a polynomial with integer coefficients, not all zero, whose coefficients' absolute values and degree add up to *some* positive integer N, A is the set of all algebraic numbers. Moreover, as the union of an at most countable number of at most countable sets, the Corollary to Theorem 2.12 implies that A is at most countable. Additionally, since the set of solutions to $a_0 z + a_1 = 0$ for $a_0, a_1 \in \mathbb{Z}$, $a_0 \neq 0$ is both a subset of the algebraic numbers and equal to $\mathbb{Q}$ (a countable set), A is at least countable. Therefore, A is countable, as desired. □

3. Prove that there exist real numbers which are not algebraic.

Proof. Suppose for the sake of contradiction that every real number is algebraic. Then if A is the set of all complex algebraic numbers, $\mathbb{R} \subset A$. Thus, since $\mathbb{R}$ is infinite and A is countable (by Problem 2.2), Theorem 2.8 implies that $\mathbb{R}$ is countable, a contradiction. □

4. Is the set of all irrational real numbers countable?

Proof. No.

Suppose for the sake of contradiction that $\mathbb{R} \setminus \mathbb{Q}$ is countable. Then since $\mathbb{R} \setminus \mathbb{Q}$ and $\mathbb{Q}$ are at most countable, the Corollary to Theorem 2.12 implies that $(\mathbb{R} \setminus \mathbb{Q}) \cup \mathbb{Q} = \mathbb{R}$ is at most countable, contradicting the fact that $\mathbb{R}$ is uncountable. □

5. Construct a bounded set of real numbers with exactly three limit points.

Proof. Let $A = \bigcup_{i=0}^{2} \{1/n + i : n \in \mathbb{N}\}$. Then A has limit points at $0, 1, 2$ and nowhere else. □

6. Let E' be the set of all limit points of a set E. Prove that E' is closed. Prove that E and $\bar{E}$ have the same limit points (recall that $\bar{E} = E \cup E'$). Do E and E' always have the same limit points?

Proof. To prove that E' is closed, it will suffice to show that it contains all of its limit points. Let p be an arbitrary limit point of E'. To show that $p \in E'$, it will suffice to verify that p is a limit point of E, i.e., that every neighborhood of p intersects E at a point other than p. Let $N_r(p)$ be an arbitrary neighborhood of p. Since p is a limit point of E', $N_r(p) \cap E'$ is infinite (2.20). Thus, we can choose a point $x \in N_r(p) \cap E'$ such that $x \neq p$. It follows that $x \in E'$, so it must be that every neighborhood of x has infinite intersection with E (2.20). In particular, since $N_r(p)$ is open and $x \in N_r(p)$, x is an interior point of $N_r(p)$, so we can choose a neighborhood N of x such that $N \subset N_r(p)$. The last two statements combined imply that $N \cap E$ is infinite. In particular, since $N \cap E \subset N \subset N_r(p)$, there exist infinitely many points of E in $N_r(p)$; choosing any one of these that is not equal to p completes the proof.

To prove that E and $\bar{E}$ have the same limit points, it will suffice to show that every limit point of E is a limit point of $\bar{E}$ and that every limit point of $\bar{E}$ is a limit point of E. The latter was accomplished by the above. Thus, let p be an arbitrary limit point of E. To prove that p is a limit point of $\bar{E}$, it will suffice to show that every neighborhood of p intersects $\bar{E}$ at some point other than p. Consider an arbitrary neighborhood $N_r(p)$ of p. Since p is a limit point of E, $N_r(p) \cap (E \setminus \{p\}) \neq \emptyset$. Therefore, we have that

$$\begin{aligned}
N_r(p) \cap (\bar{E} \setminus \{p\}) &= N_r(p) \cap [(E \cup E') \setminus \{p\}] \\
&= N_r(p) \cap [(E \setminus \{p\}) \cup (E' \setminus \{p\})] \\
&= [N_r(p) \cap (E \setminus \{p\})] \cup [N_r(p) \cap (E' \setminus \{p\})] \\
&\supset N_r(p) \cap (E \setminus \{p\}) \\
&\neq \emptyset
\end{aligned}$$

as desired.

No, E and E' do not always have the same limit points. Let $E = \{1/n : n \in \mathbb{N}\}$. Then $E' = \{0\}$, but since E' is finite, $E'' = \emptyset$. $\square$

7. Let $A_1, A_2, \dots$ be subsets of a metric space.

(a) If $B_n = \bigcup_{i=1}^n A_i$, prove that $\bar{B}_n = \bigcup_{i=1}^n \bar{A}_i$ for $n = 1, 2, 3, \dots$.

Proof. Let $n \in \mathbb{N}$ be arbitrary.

Suppose first that $x \in \bar{B}_n$. We divide into two cases ($x \in B_n$ and $x \in B_n'$). If $x \in B_n$, then $x \in A_i$ for some $i = 1, \dots, n$. It follows that $x \in A_i \cup A_i' = \bar{A}_i \subset \bigcup_{i=1}^n \bar{A}_i$, as desired. On the other hand, if $x \in B_n'$, then $N_r(x) \cap (B_n \setminus \{p\}) \neq \emptyset$ for every $r > 0$. Now suppose for the sake of contradiction that $x \notin A_i'$ for any $i = 1, \dots, n$. Then there exist neighborhoods $N_{r_1}(x), \dots, N_{r_n}(x)$ of x such that no $N_{r_i}(x)$ contains a point of A_i other than p. Let $0 < r_j \leq r_i$ for each $i = 1, \dots, n$. It follows that

$$\begin{aligned}
\emptyset &= \bigcup_{i=1}^n N_{r_j}(x) \cap (A_i \setminus \{p\}) \\
&= N_{r_j}(x) \cap \left[\bigcup_{i=1}^n (A_i \setminus \{p\})\right] \\
&= N_{r_j}(x) \cap \left[\left(\bigcup_{i=1}^n A_i\right) \setminus \{p\}\right] \\
&= N_{r_j}(x) \cap [B_n \setminus \{p\}]
\end{aligned}$$

a contradiction. Therefore, $x \in A_i'$ for some $i = 1, \dots, n$. It follows that $x \in A_i \cup A_i' = \bar{A}_i \subset \bigcup_{i=1}^n \bar{A}_i$, as desired.

Now suppose that $x \in \bigcup_{i=1}^n \bar{A}_i$. Then $x \in \bar{A}_i$ for some $i = 1, \dots, n$. We divide into two cases ($x \in A_i$ and $x \in A_i'$). If $x \in A_i$, then $x \in \bigcup_{i=1}^n A_i = B_n \subset B_n \cup B_n' = \bar{B}_n$, as desired. On the other hand, if $x \in A_i'$, then every neighborhood of x contains a point $q \neq x$ of A_i. But since $A_i \subset \bigcup_{i=1}^n A_i = B_n$, it follows that every neighborhood of x contains a point $q \neq x$ of B_n. Thus, $x \in B_n' \subset B_n \cup B_n' = \bar{B}_n$, as desired. ☐

(b) If $B = \bigcup_{i=1}^\infty A_i$, prove that $\bar{B} \supset \bigcup_{i=1}^\infty \bar{A}_i$. Show, by an example, that this inclusion can be proper.

Proof. Let $x \in \bigcup_{i=1}^\infty \bar{A}_i$ be arbitrary. Then $x \in \bar{A}_i$ for some i. We divide into two cases ($x \in A_i$ and $x \in A_i'$). If $x \in A_i$, then $x \in \bigcup_{i=1}^\infty A_i = B \subset B \cup B' = \bar{B}$, as desired. On the other hand, if $x \in A_i'$, then every neighborhood of x contains a point $q \neq x$ of A_i. But since $A_i \subset \bigcup_{i=1}^\infty A_i = B$, it follows that every neighborhood of x contains a point $q \neq x$ of B. Thus, $x \in B' \subset B \cup B' = \bar{B}$, as desired.

Define the family of sets $\{A_n\}$ by $A_n = \{1/n\}$ for each $n \in \mathbb{N}$. Then since each A_n is finite, each $\bar{A}_n = \emptyset$, so $\bigcup_{i=1}^\infty \bar{A}_i = \emptyset$. However, $B = \bigcup_{i=1}^\infty A_i$ has zero as a limit point, so

$$\bar{B} \supset \{0\} \supsetneq \emptyset = \bigcup_{i=1}^\infty \bar{A}_i$$

as desired. ☐

8. Is every point of every open set $E \subset \mathbb{R}^2$ a limit point of E? Answer the same question for closed sets in $\mathbb{R}^2$.

Proof. Yes, every point of every open set $E \subset \mathbb{R}^2$ is a limit point of E. Let E be an arbitrary open subset of $\mathbb{R}^2$. Let $x \in E$ be arbitrary. Since $x \in E$ open, x is an interior point of E, meaning that there exists $N_r(x) \subset E$. Now to prove that x is a limit point of E, it will suffice to show that every neighborhood of x contains a point $q \neq x$ of E. Let $N_s(x)$ be an arbitrary neighborhood of x. If $x = (x_1, x_2)$ and $m = \min(r, s)$, choose $q = (x_1 + m/2, x_2 + m/2)$. Since $r, s > 0$ by definition, $q \neq x$. Additionally,

$$\begin{aligned}
|q - x|^2 &= (x_1 + m/2 - x_1)^2 + (x_2 + m/2 - x_2)^2 \\
&= m^2/2 \\
&< m^2
\end{aligned}$$

Taking square roots reveals that $|q - x| < r$ and $|q - x| < s$. It follows that $q \in N_r(x) \subset E$ and $q \in N_s(x)$, as desired.

No, every point of every closed set $E \subset \mathbb{R}^2$ is not a limit point of E. Let E be a nonempty finite set. Then by the table on Rudin (1976, p. 33), E is closed but not perfect, implying that E is a closed set not every point of which is a limit point of it (in fact, the fact that not every point of every closed set is a limit point of E is the whole motivation for defining perfect sets!). ☐

9. Let E° denote the set of all interior points of a set E (see Definition 2.18e; E° is called the **interior** of E).

(a) Prove that E° is always open.

Proof. Let $x \in E^\circ$ be arbitrary. Then since x is an interior point of E, there exists a neighborhood $N(x)$ of x such that $N(x) \subset E$. By Theorem 2.19, $N(x)$ is open. It follows from Theorem 2.24 that $\bigcup_{x \in E^\circ} N(x)$ is open. We now prove that $E^\circ = \bigcup_{x \in E^\circ} N(x)$. The inclusion in one direction is obvious. In the other, let $y \in \bigcup_{x \in E^\circ} N(x)$. Then $y \in N(x)$ for some x. It follows since each $N(x)$ is open that there exists a neighborhood N of y such that $N \subset N(x)$. But since $N(x) \subset E$ by definition, we have both that $y \in E$ and that $N \subset E$. Thus, y is an interior point of E, so $y \in E^\circ$, as desired. ☐

(b) Prove that E is open if and only if $E^\circ = E$.

Proof. Suppose first that E is open. Let $x \in E^\circ$ be arbitrary. Then since x is an interior point of E, x is naturally a point of E. On the other hand, let $x \in E$. Then since E is open, x is an interior point of E, so $x \in E^\circ$, as desired.

Now suppose that $E^\circ = E$. Then since E° is open by part (a), E is open. $\square$

(c) If $G \subset E$ and G is open, prove that $G \subset E^\circ$.

Proof. Let $x \in G$ be arbitrary. Since G is open, there exists a neighborhood N of x such that $N \subset G$. But since $G \subset E$, $N \subset E$. Thus, x is an interior point of E, so $x \in E^\circ$, as desired. $\square$

(d) Prove that the complement of E° is the closure of the complement of E.

Proof. Let $x \in (E^\circ)^c$. Then $x \notin E^\circ$. We divide into two cases ($x \notin E$ and $x \in E$). If $x \notin E$, then $x \in E^c$. It follows that $x \in E^c \cup (E^c)' = \overline{E^c}$, as desired. On the other hand, if $x \in E$ (but $x \notin E^\circ$), then there exists no neighborhood of x that is a subset of E. In other words, every neighborhood of x contains some point of E^c. This combined with the fact that $x \notin E^c$ implies that $x \in (E^c)'$. Therefore, $x \in E^c \cup (E^c)' = \overline{E^c}$, as desired.

Let $x \in \overline{E^c}$. We divide into two cases ($x \in E^c$ and $x \in (E^c)'$). If $x \in E^c$, then $x \notin E$. It follows that $x \notin E^\circ \subset E$. Therefore, $x \in (E^\circ)^c$, as desired. On the other hand, if $x \in (E^c)'$, then every neighborhood of x contains a point of E^c. This combined with the fact that $x \in E$ ($x \notin E^c$ in this case) implies that no neighborhood N of x exists such that $N \subset E$. Therefore, x is not an interior point of E, i.e., $x \notin E^\circ$; it follows that $x \in (E^\circ)^c$, as desired. $\square$

(e) Do E and $\bar{E}$ always have the same interiors?

Proof. No.

Consider $\mathbb{Q} \subset \mathbb{R}$. Since $\mathbb{Q}$ is disconnected at every point, $\mathbb{Q}^\circ = \emptyset$ but $(\bar{\mathbb{Q}})^\circ = \mathbb{R}^\circ = \mathbb{R}$. $\square$

(f) Do E and E° always have the same closures?

Proof. No.

Consider $\mathbb{Q} \subset \mathbb{R}$. As before, we have that $\bar{\mathbb{Q}} = \mathbb{R}$ while $\bar{\mathbb{Q}^\circ} = \bar{\emptyset} = \emptyset$. $\square$

10. Let X be an infinite set. For $p \in X$ and $q \in X$, define

$$d(p,q) = \begin{cases} 1 & p \neq q \\ 0 & p = q \end{cases}$$

Prove that this is a metric. Which subsets of the resulting metric space are open? Which are closed? Which are compact?

Proof. To prove that d is a metric, it will suffice to show that $d(p,q) > 0$ if $p \neq q$, $d(p,p) = 0$, $d(p,q) = d(q,p)$, and $d(p,q) \leq d(p,r) + d(r,q)$ for any $r \in X$. Let's begin. Let $p \neq q$. Then by the definition of d, $d(p,q) = 1 > 0$, as desired. Let $p \in X$. Then by the definition of d, $d(p,p) = 0$, as desired. Let $p, q \in X$. We divide into two cases ($p = q$ and $p \neq q$). If $p = q$, then $d(p,q) = 0 = d(q,p)$. If $p \neq q$, then $d(p,q) = 1 = d(q,p)$, as desired. Let $p, q, r \in X$. We divide into two cases ($p = q$ and $p \neq q$). If $p = q$, then $d(p,q) = 0$ must be less than the sum of two numbers that are either 0 or 1. If $p \neq q$, then $d(p,q) = 1$. However, since r cannot equal the distinct p and q, at least on of $d(p,r)$ and $d(r,q)$ equals 1, so the inequality holds here, too, as desired.

Every subset is open. Let $E \subset X$, and let $x \in E$. Then by the definition of d, $N_1(x) = \{y \in X : d(y,x) < 1\} = \{x\} \subset E$. Thus, every point of E is an interior point, as desired.

Every subset is closed. Let $E \subset X$. By the previous result, E^c is open. Thus, by Theorem 2.23, E is closed.

Only finite sets are compact. We know that every finite set is compact (choose an open cover $\{G_\alpha\}$ of E finite; map every $x \in E$ to some G_α that contains it; choose the range of this map as the finite subcover). If E is infinite, however, choose the open cover $\{\{x\}\}_{x \in E}$. We know that all of these sets are open (because every set is open). Additionally, since each one only contains one element of E, we need all infinitely many of them to cover E. Thus, this infinite E is not compact. $\square$

11. For $x \in \mathbb{R}^1$ and $y \in \mathbb{R}^1$, define

$$\begin{aligned} d_1(x,y) &= (x-y)^2 \\ d_2(x,y) &= \sqrt{|x-y|} \\ d_3(x,y) &= |x^2-y^2| \\ d_4(x,y) &= |x-2y| \\ d_5(x,y) &= \frac{|x-y|}{1+|x-y|} \end{aligned}$$

Determine, for each of these, whether it is a metric or not.

Proof. d_1 is not a metric. Let $x=2$, $y=0$, $z=1$. Then

$$d_1(2,0) = (2-0)^2 = 4 > 2 = (2-1)^2 + (1-0)^2 = d_1(2,1) + d_1(1,0)$$

so d_1 does not obey the triangle inequality.

d_2 is a metric. If $x \neq y$, then $x-y \neq 0$, so $d_2(x,y) = \sqrt{|x-y|} > 0$, as desired. For each x, $d_2(x,x) = \sqrt{|x-x|} = \sqrt{0} = 0$, as desired. For all x,y, $d_2(x,y) = \sqrt{|x-y|} = \sqrt{|y-x|} = d_2(y,x)$, as desired. For all x,y,z,

$$\begin{aligned} d_2(x,y) &= \sqrt{|x-y|} \\ &\leq \sqrt{|x-z|+|z-y|} \\ &\leq \sqrt{|x-z|} + \sqrt{|z-y|} \\ &= d_2(x,z) + d_2(z,y) \end{aligned}$$

as desired.

d_3 is not a metric. Let $x=1$, $y=-1$. Then $x \neq y$, but

$$d_3(1,-1) = |1^2 - (-1)^2| = 0$$

d_4 is not a metric. Let $x=2$, $y=1$. Then $x \neq y$, but

$$d_4(2,1) = |2-2(1)| = 0$$

d_5 is a metric. If $x \neq y$, then $x-y \neq 0$, so $d_5(x,y) = |x-y|/(1+|x-y|) > 0$, as desired. For each x, $d_5(x,x) = |x-x|/(1+|x-x|) = 0$, as desired. For all x,y, $d_5(x,y) = |x-y|/(1+|x-y|) = |y-x|/(1+|y-x|) = d(y,x)$. For all x,y,z,

$$\begin{aligned} d(x,y) &= \frac{|x-y|}{1+|x-y|} \\ &\leq \frac{|x-z|+|z-y|}{1+|x-z|+|z-y|} \\ &= \frac{|x-z|}{1+|x-z|+|z-y|} + \frac{|z-y|}{1+|x-z|+|z-y|} \\ &\leq \frac{|x-z|}{1+|x-z|} + \frac{|z-y|}{1+|z-y|} \\ &= d(x,z) + d(z,y) \end{aligned}$$

as desired. □

12. Let $K \subset \mathbb{R}^1$ consist of 0 and the numbers $1/n$ for $n = 1,2,3,\ldots$. Prove that K is compact directly from the definition (without using the Heine-Borel theorem).

Proof. Let $\{G_\alpha\}$ be an arbitrary open cover of K. Then $0 \in G_\alpha$ for some α. Since G_α is open, 0 is an interior point of it, so there exists a neighborhood $N_r(0)$ such that $N_r(0) \subset G_\alpha$. Since $r > 0$ by definition, if we let $x = r$ and $y = 1$, the Archimedean property implies there exists a positive integer m such that $mr > 1$. It follows that $1/m < r$, so every $1/n$ such that $n \geq m$ is an element of $N_r(0) \subset G_\alpha$. Since G_α contains 0 and infinitely many of the $1/n$, let this G_α be part of our finite subcover. For the remaining entries in our finite subcover, choose for each of the finitely many $1/n$ such that $n < m$ a G_β that contains it. ☐

13. Construct a compact set of real numbers whose limit points form a countable set.

Proof. Consider the family of sets $\{K_i\}$ defined by

$$K_i = \{1/i\} \cup \{1/i + 1/n : n \in \mathbb{N}\}$$

for each $i \in \mathbb{N}$ and $i = +\infty$. Let

$$K = \bigcup_{i=1}^{+\infty} K_i$$

K is bounded with lower bound $0 \in K_\infty$ and upper bound $2 = 1/1 + 1/1 \in K_1$. Additionally, K is closed with limit points $K' = K_\infty$. Thus, if we define $f : \mathbb{N} \to K'$ by

$$f(n) = \begin{cases} 0 & n = 1 \\ \frac{1}{n-1} & n > 1 \end{cases}$$

we will have a bijection between the natural number and K', proving that K' is countable, as desired. ☐

14. Give an example of an open cover of the segment $(0, 1)$ which has no finite subcover.

Proof. Choose $\{G_i\}_{i=3}^{\infty}$ defined by

$$G_i = \left(\frac{1}{i}, \frac{1}{i-2} \right)$$

Every segment is open in $\mathbb{R}$. Additionally, $\{G_i\}$ is a cover since if $x \in (0, 1)$, then we can modify the Archimedean property to choose the smallest integer n such that $1/n < x$. It follows that $x \leq \frac{1}{n-1} < \frac{1}{n-2}$, so $x \in (1/n, 1/(n-2))$, as desired. Lastly, $\{G_i\}$ has no finite subcover: if it did, we could use the betweeness of the reals to choose an $x < 1/i$ where $(1/i, 1/(i-2))$ is the smallest segment in the finite subcover. It would follow that $x \in (0, 1)$ but x is not an element of any set in the cover, a contradiction. ☐

15. Show that Theorem 2.36 and its Corollary become false (in $\mathbb{R}^1$, for example) if the word "compact" is replaced by "closed" or by "bounded."

Proof. Suppose first that "compact" is replaced by "closed." Consider the collection of sets $\{K_n\}_{n=1}^{\infty}$ defined by

$$K_n = n\mathbb{N}$$

for each n, where by $n\mathbb{N}$ we mean all the natural number multiples of n (e.g., $3\mathbb{N} = \{3, 6, 9, ...\}$). Clearly any finite collection of these sets will intersect at the least common multiple of the relevant n's. However, the intersection of all such sets will be the empty set since for any possible natural number n in the intersection, $n \notin (n + 1)\mathbb{N} = K_{n+1}$.

Now suppose that "compact" is replaced by "bounded." Consider the collection of sets $\{K_n\}_{n=1}^{\infty}$ defined by

$$K_n = (0, 1/n)$$

for each n. This family of sets satisfies the constraints of both the modified Theorem 2.36 and its Corollary. However, $\bigcap_{n=1}^{\infty} K_n = \emptyset$ since by the Archimedean principle, we can always find a $1/n$ smaller than any x in any of the sets, and thus a set in the intersection that does not contain said x. ☐

7 Basic Topology II

From Rudin (1976).

Chapter 2

11/15: **16.** Regard $\mathbb{Q}$, the set of all rational numbers, as a metric space, with $d(p,q) = |p-q|$. Let E be the set of all $p \in \mathbb{Q}$ such that $2 < p^2 < 3$. Show that E is closed and bounded in $\mathbb{Q}$, but that E is not compact. Is E open in $\mathbb{Q}$?

Proof. In this proof, we treat the set of all $x \in E$ such that $x > 0$ and $x < 0$ separately, knowing that the union of two (a finite number of) closed, bounded, or compact sets will be closed, bounded, or compact. We will directly treat the $x > 0$ case; the proof of the other case is symmetric (this means that when we way "E" herein, we are referring to the subset of E containing only the positive elements of E). Let's begin.

Suppose for the sake of contradiction that there exists $x \in E'$ such that $x \notin E$. Since $x \notin E$, $2 \not< x^2 \not< 3$, i.e., $x^2 \leq 2$ or $x^2 \geq 3$. Moreover, since x is rational and no rational number has square equal to 2 or 3, $x^2 < 2$ or $x^2 > 3$. We now divide into two cases. Suppose $x^2 < 2$. Let $y = 2(x+1)/(x+2)$. Then $y^2 - 2 = 2(x^2-2)/(x+2)^2 < 0$ since $x^2 - 2 < 0$, so $y^2 < 2$. It follows that y is a lower bound on E since for any $z \in E$, $y^2 < 2 < z^2$ so $y < z$. Now consider the neighborhood $N_r(x)$ where $r = d(y,x)$. Since $y > x$, y is clearly an upper bound on this neighborhood as defined (if $z > y$, then $d(z,x) > d(y,x)$). Thus, if $p \in E$, $p^2 > 2 > y^2$ so $p > y$, so p cannot be $< y$ as would be necessary for it to be in $N_r(x)$. Thus, having found a neighborhood of x having empty intersection with E, we know that $x \notin E'$, a contradiction. Now suppose $x^2 > 3$. Then consider $y = 3(x+1)/(x+3)$ and follow through a symmetric argument to a contradiction.

To prove that E is bounded, let $x = 0$ and $y = 2$. Suppose $p \in E$. Then $x^2 = 0 < p^2 < 3 < 4 = y^2$. It follows that $x < p < y$. Thus, E is bounded below by $x = 0$ and bounded above by $y = 2$.

Define an open cover $\{G_p\}$ of E by

$$G_p = \left(\frac{2(p+1)}{p+2}, \frac{3(p+1)}{p+3} \right)$$

for each $p \in E$. Clearly each G_p is open as a segment and $p \in G_p$ for each $p \in E$, so $\{G_p\}$ is an open cover of E. Now suppose that $\{G_{p_n}\}$ is a finite open cover of E. Choose $p = \max p_n$. Then $2 < 3(p+1)/(p+3) < 3$, but $3(p+1)/(p+3) \notin G_{p_n}$ for any n, a contradiction. Thus, G is not compact, as desired.

E is open in $\mathbb{Q}$ since for any point $p \in E$, we may choose the segment $G_p \subset E$ containing p and then choose a neighborhood subset of G_p. □

17. Let E be the set of all $x \in [0,1]$ whose decimal expansion contains only the digits 4 and 7. Is E countable? Is E dense in $[0,1]$? Is E compact? Is E perfect?

Proof. No, E is not countable. Suppose for the sake of contradiction that E is countable. Then every infinite subset $F \subset E$ is countable. Let $F = \{x_1, x_2, \dots\}$ be such a subset. Define $x \in E$ by the rule that if the n^{th} digit in the decimal expansion of x_n is 4, we let the n^{th} digit in the decimal expansion of x be 7 (and vice versa). Thus, x differs from every element of F by at least one decimal point, so $x \notin F$. It follows that every countable subset of E is a proper subset of E, meaning that E itself must be uncountable (for otherwise we would have $E \subsetneq E$).

No, E is not dense in $[0,1]$. Suppose for the sake of contradiction that E is dense in $[0,1]$. Then $\inf E = 0$. It follows that there exists $x \in E$ such that $0 \leq x < 0.1$. But since $x < 0.1$, the first digit in its decimal expansion is necessarily 0 (i.e., not 4 or 7), contradicting the hypothesis that $x \in E$.

Yes, E is compact. To prove this, we will show that E is closed and then invoke Theorem 2.35 since $E \subset [0,1]$ compact. Suppose for the sake of contradiction that there exists a $p \in E'$ such that $p \notin E$. Let

$$p = \sum_{k=1}^{\infty} \frac{n_k}{10^k}$$

be the decimal expansion of p (we need not consider a n_0 term since $p \in [0,1]$ if p is a limit point of E). Then since $p \notin E$, there exists an $n_k \neq 4, 7$. Let N be the smallest positive integer such that $n_N \neq 4, 7$. Now let $q \in E$ be arbitrary, with decimal expansion

$$q = \sum_{k=1}^{\infty} \frac{m_k}{10^k}$$

It follows that

$$|q - p| = \left| \sum_{k=N}^{\infty} \frac{m_k}{10^k} - \sum_{k=N}^{\infty} \frac{n_k}{10^k} \right|$$

We sum from N because the first $N-1$ equal terms cancel.

$$\begin{aligned}
&= \left| \sum_{k=N}^{\infty} \frac{1}{10^k}(m_k - n_k) \right| \\
&= \left| \frac{1}{10^N}(m_N - n_N) - \sum_{k=N+1}^{\infty} \frac{1}{10^k}(n_k - m_k) \right| \\
&\geq \left| \left| \frac{1}{10^N}(m_N - n_N) \right| - \left| \sum_{k=N+1}^{\infty} \frac{1}{10^k}(m_k - n_k) \right| \right| \\
&= \left| \frac{1}{10^N}(m_N - n_N) \right| - \left| \sum_{k=N+1}^{\infty} \frac{1}{10^k}(m_k - n_k) \right|
\end{aligned}$$

The above equality holds since the nonzero difference of the N^{th} terms of the decimal expansion will necessarily be greater than the sum of the rest of the terms, each of which is bounded above by $7/10^k$ (given in the case that $m_k = 7$ and $n_k = 0$).

$$\begin{aligned}
&= \frac{1}{10^N}|m_N - n_N| - \left| \sum_{k=N+1}^{\infty} \frac{1}{10^k}(m_k - n_k) \right| \\
&\geq \frac{1}{10^N}(1) - \sum_{k=N+1}^{\infty} \frac{1}{10^k}|m_k - n_k| \\
&\geq \frac{1}{10^N} - \sum_{k=N+1}^{\infty} \frac{1}{10^k}(7) \\
&= \frac{1}{10^N} - \frac{7}{10^{N+1}} \sum_{k=0}^{\infty} \left(\frac{1}{10}\right)^k \\
&= \frac{1}{10^N} - \frac{7}{10^{N+1}} \cdot \frac{1}{1 - 1/10} && \text{Theorem 3.26} \\
&= \frac{2}{9 \cdot 10^N}
\end{aligned}$$

Thus, $2/(9 \cdot 10^N)$ is a lower bound on $|q - p|$. However, since $p \in E'$, we can choose $N_{1/10^N}(p)$ and know that there exists a $q \in E$ such that $q \in N_{1/10^N}(p)$. But then this q gives

$$|q - p| < \frac{1}{10^N} < \frac{2}{9 \cdot 10^N} \leq |q - p|$$

a contradiction.

Yes, E is perfect. To prove this, it will suffice to show that every $x \in E$ is a limit point of E. Let $p \in E$ be arbitrary, and let $N_r(p)$ be any neighborhood of p. By the Archimedean property, find an $m \in \mathbb{N}$ such that $mr > 1$. Now round m up to the next multiple of 10, i.e., find n such that $10^n \geq m$. It follows that $1/10^n < r$. Let q be the number with decimal expansion identical to p except at the $(n+1)^{\text{th}}$ slot, where whatever's there (4 or 7) is flipped (to 7 or 4). It follows that

$$|q - p| = \frac{7-4}{10^{n+1}} < \frac{1}{10^n} < r$$

so $q \neq p$ satisfies $q \in N_r(p)$, as desired. $\square$

18. Is there a nonempty perfect set in $\mathbb{R}^1$ which contains no rational number?

Proof. Yes.

Consider the set E from Exercise 2.17. Define F by

$$F = \{x + 0.101001000100001... : x \in E\}$$

Since E is nonempty and perfect, F (as nothing but a rigid translation of E) will also be nonempty and perfect. Additionally, every element of F is irrational since it will be a sequence of 4's and 7's interrupted by 5's and 8's at nonrepeating intervals. $\square$

19. (a) If A and B are disjoint closed sets in some metric space X, prove that they are separated.

Proof. Since A, B are closed, consecutive applications of Theorem 2.27b imply that $A = \bar{A}$ and $B = \bar{B}$. Therefore, since A, B are disjoint,

$$\emptyset = A \cap B = A \cap \bar{B} \qquad\qquad \emptyset = A \cap B = \bar{A} \cap B$$

as desired. $\square$

(b) Prove the same for disjoint open sets.

Proof. Let A, B be disjoint open sets. Suppose for the sake of contradiction that $\bar{A} \cap B \neq \emptyset$. Then there exists $x \in (A \cup A') \cap B = A' \cap B$. Since $x \in B$ open, x is an interior point of B, so there exists a neighborhood $N_r(x) \subset B$. But since $N_r(x)$ is a neighborhood of x and $x \in A'$, $N_r(x) \cap (A \setminus \{x\}) \neq \emptyset$, so there exists a point $y \in A$ such that $y \in N_r(x)$. It follows since $N_r(x) \subset B$ that $y \in B$. Therefore, since $y \in A$ and $y \in B$, $y \in A \cap B \neq \emptyset$, a contradiction. $\square$

(c) Fix $p \in X$ and $\delta > 0$. Define A to be the set of all $q \in X$ for which $d(p, q) < \delta$, and define B similarly with $>$ in place of $<$. Prove that A and B are separated.

Proof. We have by the definition of A that $A = N_\delta(p)$. Thus, Theorem 2.19 implies that A is open. Additionally, $B^c = \{q \in X : d(p, q) \leq \delta\}$ is closed. Thus, Theorem 2.23 implies that B is open. Applying part (b) therefore gives the desired result. $\square$

(d) Prove that every connected metric space with at least two points is uncountable. (Hint: Use (c).)

Proof. Let X be a connected metric space containing two distinct points x, y. Suppose for the sake of contradiction that X is at most countable. Consider the set

$$\Delta = \{d(p, x) : p \in X\}$$

Clearly $\tilde{d} : X \to \Delta$ defined by $\tilde{d}(x) = d(p, x)$ is a surjection, so Δ is an at most countable collection of nonnegative real numbers. It follows by an argument analogous to that used in Problem 2.4 that $\mathbb{R}^+ \setminus \Delta$ is uncountable. Thus, we may choose a $\delta \in \mathbb{R}^+ \setminus \Delta$. Clearly $\delta > 0$ and since $\delta \notin \Delta$, there exists no $p \in X$ such that $d(p, x) = \delta$. Consequently, we may define A to be the set of all $p \in X$ for which $d(p, x) < \delta$ and define B similarly with $>$ in place of $<$ and know that $A \cup B = X$. Additionally, it follows by (c) that A, B are separated, so X is not connected, a contradiction. Therefore, X is uncountable, as desired. $\square$

20. Are closures and interiors of connected sets always connected? (Hint: Look at subsets of $\mathbb{R}^2$.)

Proof. Yes, the closures of connected sets are always connected.

Let X be a connected set. Suppose for the sake of contradiction that $\bar{X}$ is disconnected. Then there exist separated sets A, B such that $A \cup B = \bar{X}$. We now seek to construct from A, B separated sets whose union is equal to X. Choose $A \setminus (X' \setminus X), B \setminus (X' \setminus X)$. We know that the union of these two sets is equal to X since

$$\begin{aligned} A \setminus (X' \setminus X) \cup B \setminus (X' \setminus X) &= (A \cup B) \setminus (X' \setminus X) \\ &= \bar{X} \setminus (X' \setminus X) \\ &= (X \cup X') \setminus (X' \setminus X) \\ &= X \end{aligned}$$

We now seek to prove that $A \setminus (X' \setminus X), B \setminus (X' \setminus X)$ are nonempty. Suppose first for the sake of contradiction that $A \setminus (X' \setminus X) = \emptyset$. It follows that

$$\begin{aligned} A \cup B &= \bar{X} \\ A \cup B &= X \cup X' \\ (A \cup B) \setminus (X' \setminus X) &= (X \cup X') \setminus (X' \setminus X) \\ (A \setminus (X' \setminus X)) \cup (B \setminus (X' \setminus X)) &= X \\ B \setminus (X' \setminus X) &= X \\ B &\supset X \end{aligned}$$

Consequently, $\bar{B} \supset \bar{X}$. But then since $A \subset \bar{X}$ and $A \cap \bar{B} = \emptyset$, $A = \emptyset$, a contradiction. A symmetric argument proves that $B \setminus (X' \setminus X)$ is nonempty. Lastly, we have that

$$(A \setminus (X' \setminus X)) \cap \overline{B \setminus (X' \setminus X)} \subset A \cap \bar{B} = \emptyset$$

and symmetrically for the other requirement, so X is disconnected, a contradiction.

Interiors of connected sets are not always connected.

Consider the set $X = \overline{N_1(1,0)} \cup \overline{N_1(-1,0)} \subset \mathbb{R}^2$. X is connected since both closed ball components are connected and the two components have nonempty intersection. Moreover, $X^\circ = N_1(1,0) \cup N_1(-1,0)$ is disconnected: Choose $A = N_1(1,0)$ and $B = N_1(-1,0)$. Then A, B are nonempty. Additionally, $A \cap \bar{B} = \emptyset$ since every point of A has x-component greater than zero and every point of $\bar{B}$ has x component less than or equal to zero. ☐

21. Let A and B be separated subsets of some $\mathbb{R}^k$, suppose $\mathbf{a} \in A$ and $\mathbf{b} \in B$, and define

$$\mathbf{p}(t) = (1-t)\mathbf{a} + t\mathbf{b}$$

for all $t \in \mathbb{R}^1$. Let $A_0 = \mathbf{p}^{-1}(A)$, $B_0 = \mathbf{p}^{-1}(B)$.

(a) Prove that A_0 and B_0 are separated subsets of $\mathbb{R}^1$.

Proof. To prove that A_0, B_0 are separated subsets of $\mathbb{R}^1$, it will suffice to show that $A_0 \cap \bar{B}_0 = \emptyset$ and $\bar{A}_0 \cap B = \emptyset$. To begin, suppose for the sake of contradiction that $A_0 \cap \bar{B}_0 \neq \emptyset$. Then there exists t such that $t \in A_0$ and $t \in \bar{B}_0$. Since $t \in A_0$, $\mathbf{p}(t) \in A$. We now divide into two cases ($t \in B_0$ and $t \in B_0'$). If $t \in B_0$, then $\mathbf{p}(t) \in B$. But then

$$A \cap \bar{B} \supset A \cap B \supset \{\mathbf{p}(t)\} \neq \emptyset$$

so A and B are not separated, a contradiction. On the other hand, if $t \in B_0'$, we can show that $\mathbf{p}(t) \in B'$: Let $N_r[\mathbf{p}(t)]$ be an arbitrary neighborhood of $\mathbf{p}(t)$. Since $t \in B_0'$, we know that $N_{r/\|\mathbf{a}-\mathbf{b}\|}(t) \cap (B_0 \setminus \{t\}) \neq \emptyset$ (note that $\|\mathbf{a}-\mathbf{b}\| \neq 0$ since $\mathbf{a} \neq \mathbf{b}$ as members of separated [hence disjoint] sets). Thus, choose $s \in N_{r/\|\mathbf{a}-\mathbf{b}\|}(t) \cap (B_0 \setminus \{t\})$. Consequently, $|s-t| < r/\|\mathbf{a}-\mathbf{b}\|$, $s \in B_0$, and $s \neq t$. It follows from the first condition that

$$\begin{aligned}\|\mathbf{p}(t) - \mathbf{p}(s)\| &= \|[(1-t)\mathbf{a} + t\mathbf{b}] - [(1-s)\mathbf{a} + s\mathbf{b}]\| \\ &= \|(s-t)\mathbf{a} - (s-t)\mathbf{b}\| \\ &= |s-t| \cdot \|\mathbf{a}-\mathbf{b}\| \\ &< \frac{r}{\|\mathbf{a}-\mathbf{b}\|} \cdot \|\mathbf{a}-\mathbf{b}\| \\ &= r\end{aligned}$$

so $\mathbf{p}(s) \in N_r[\mathbf{p}(t)]$. It follows from the second condition that $\mathbf{p}(s) \in B$. It follows from the third condition that $\mathbf{p}(s) \neq \mathbf{p}(t)$. These last three results combined imply that $\mathbf{p}(s) \in N_r[\mathbf{p}(t)] \cap (B \setminus \{\mathbf{p}(t)\})$. Therefore, $\mathbf{p}(t) \in B'$, as desired. This combined with the fact that $\mathbf{p}(t) \in A$ implies that

$$A \cap \bar{B} \supset A \cap B' \supset \{\mathbf{p}(t)\} \neq \emptyset$$

so A and B are not separated, a contradiction. The proof that $\bar{A}_0 \cap B = \emptyset$ is symmetric. □

(b) Prove that there exists $t_0 \in (0,1)$ such that $\mathbf{p}(t_0) \notin A \cup B$.

Proof. Let $\tilde{A}_0 = A_0 \cap (0,1)$ and let $\tilde{B}_0 = B_0 \cap (0,1)$. Since A_0, B_0 are separated by part (a), the subsets $\tilde{A}_0, \tilde{B}_0$ of A_0, B_0, respectively, are separated, too. This means that $\tilde{A}_0 \cup \tilde{B}_0$ is disconnected. It follows by Theorem 2.47 that there exist x, y, t_0 such that $x, y \in \tilde{A}_0 \cup \tilde{B}_0$, $x < t_0 < y$, and $t_0 \notin \tilde{A}_0 \cup \tilde{B}_0$. It follows from the first condition that $0 < x < 1$ and $0 < y < 1$. This combined with the second condition implies that $0 < x < t_0 < y < 1$. This combined with the the third condition implies that $t_0 \notin A_0 \cup B_0$ (for otherwise, we would have $t_0 \in (A_0 \cup B_0) \cap (0,1) = \tilde{A}_0 \cup \tilde{B}_0$). Therefore, from the last two results combined and the definitions of A_0, B_0, we have that $t_0 \in (0,1)$ and that $\mathbf{p}(t_0) \notin A \cup B$, as desired. □

(c) Prove that every convex subset of $\mathbb{R}^k$ is connected.

Proof. Let E be a convex subset of $\mathbb{R}^k$. Then $\lambda\mathbf{x} + (1-\lambda)\mathbf{y} \in E$ for all $\mathbf{x}, \mathbf{y} \in E$ and $0 < \lambda < 1$. Now suppose for the sake of contradiction that E is disconnected. Then E is the union of two nonempty separated subsets A, B of $\mathbb{R}^k$. Since A, B are nonempty, pick $\mathbf{a} \in A$ and $\mathbf{b} \in B$. It follows by part (b) that there exists a $0 < \lambda < 1$ such that $\lambda\mathbf{a} + (1-\lambda)\mathbf{b} \notin A \cup B = E$, a contradiction. □

22. A metric space is called **separable** if it contains a countable dense subset. Show that $\mathbb{R}^k$ is separable. (Hint: Consider the set of points which only have rational coordinates.)

Proof. By the Corollary to Theorem 2.13, $\mathbb{Q}$ is countable. By Theorem 2.13,

$$\mathbb{Q}^k = \underbrace{\mathbb{Q} \times \cdots \times \mathbb{Q}}_{k \text{ times}}$$

where "$\times$" is the Cartesian product is countable.

Now all we need to do is show that $\mathbb{Q}^k$ is dense in $\mathbb{R}^k$. To do so, it will suffice to show that every point of $\mathbb{R}^k$ is a limit point of $\mathbb{Q}^k$ or a point of $\mathbb{Q}^k$, itself. Let $\mathbf{x} \in \mathbb{R}^k$ be arbitrary. Let $N_r(\mathbf{x})$ be an arbitrary neighborhood of $\mathbf{x}$. Consider the set $K = \prod_{i=1}^k [x_i - r/k, x_i + r/k]$. With the generalized Pythagorean theorem, we can show that $K \subset N_r(\mathbf{x})$. Additionally, since we can find a rational number between any two real numbers, we can always find a rational q_i between x_i and $x_i + r/k$. Thus, $\mathbf{q} = (q_1, \dots, q_k) \in \mathbb{Q}^k$, $\mathbf{q} \in K$ (since each $q_i \in [x_i - r/k, x_i + r/k]$), and $\mathbf{q} \neq \mathbf{x}$ (since each $q_i > x_i$). It follows that $\mathbf{q} \in N_r(\mathbf{x}) \cap (\mathbb{Q}^k \setminus \{\mathbf{x}\})$, so $\mathbf{x}$ is a limit point of $\mathbb{Q}^k$, as desired. □

23. A collection $\{V_\alpha\}$ of open subsets of X is said to be a **base** for X if the following is true: For every $x \in X$ and every open set $G \subset X$ such that $x \in G$, we have $x \in V_\alpha \subset G$ for some α. In other words, every open set in X is the union of a subcollection of $\{V_\alpha\}$. Prove that every separable metric space has a *countable* base. (Hint: Take all neighborhoods with rational radius and center in some countable dense subset of X.)

Proof. Let X be a separable metric space, and let E be its countable dense subset. Consider the collection $\{N_q(e)\}_{e\in E,\ q\in\mathbb{Q}^+}$. Since E is countable by definition and the infinite $\mathbb{Q}^+ \subset \mathbb{Q}$ is countable by Theorem 2.8, Theorem 2.13 implies that $\mathbb{Q}^+ \times E$ is countable. This combined with the fact that we can define a bijection $f : \{N_q(e)\}_{e\in E,\ q\in\mathbb{Q}^+} \to \mathbb{Q}^+ \times E$ by

$$f[N_q(e)] = (q, e)$$

implies that $\{N_q(e)\}_{e\in E,\ q\in\mathbb{Q}^+}$ is countable.

Now all we need to do is show that $\{N_q(e)\}_{e\in E,\ q\in\mathbb{Q}^+}$ is a base of X. First off, $\{N_q(e)\}_{e\in E,\ q\in\mathbb{Q}^+}$ is a collection of open subsets of X by the definition of each neighborhood and the fact that neighborhoods are open (Theorem 2.19). Additionally, let $x \in X$ be arbitrary and let $G \subset X$ be an arbitrary open set containing x. Since G is open, there exists a neighborhood $N_r(x)$ such that $N_r(x) \subset G$. Since E is dense in X, x is a limit point of E. Thus, there exists $e \in E$ such that $e \in N_{r/3}(x)$. Thus, choose $q \in \mathbb{Q}^+$ such that $r/3 < q < 2r/3$. It follows that

$$x \in N_{r/3}(e) \subset N_q(e) \subset N_{2r/3}(e) \subset N_r(x) \subset G$$

as desired. □

24. Let X be a metric space in which every infinite subset has a limit point. Prove that X is separable. (Hint: Fix $\delta > 0$ and pick $x_1 \in X$. Having chosen $x_1, \dots, x_j \in X$, choose $x_{j+1} \in X$, if possible, so that $d(x_i, x_{j+1}) \geq \delta$ for each $i = 1, \dots, j$. Show that this process must stop after a finite number of steps, and that X can therefore be covered by finitely many neighborhoods of radius δ. Take $\delta = 1/n$ $(n = 1, 2, 3, \dots)$ and consider the centers of the corresponding neighborhoods.)

Proof. Let $\delta > 0$ be arbitrary. Choose $x_1 \in X$. Now suppose using strong induction that we have chosen $x_1, \dots, x_j \in X$. If there exists a point $x \in X$ such that $d(x, x_i) \geq \delta$ for each $i = 1, \dots, j$, let $x_{j+1} = x$. If no such point exists, terminate the process.

First, we show that the process described above terminates after a finite number of steps. Suppose for the sake of contradiction that the set E_δ of all x_i is infinite. By the criterion on X, E_δ must then have a limit point x. It follows that $N_{\delta/2}(x)$ contains an x_i not equal to x. Similarly, it follows that $N_{d(x,x_i)}(x)$ contains an x_j not equal to x (note that by the choice of the second neighborhood, $x_i \neq x_j$ as well). But since $x_i, x_j \in N_{\delta/2}(x)$,

$$d(x_i, x_j) \leq d(x_i, x) + d(x, x_j) < \frac{\delta}{2} + \frac{\delta}{2} = \delta$$

contradicting the fact that $d(x_i, x_j) \geq \delta$ by the way they were chosen. Therefore, the set E_δ of all x_i is finite, so the process stops after a finite number of steps.

Second, we show that $\{N_\delta(x_i)\}_{x_i\in E_\delta}$ is a finite cover of X. Suppose for the sake of contradiction that there exists $x \in X$ such that $x \notin N_\delta(x_i)$ for any $x_i \in E_\delta$. Naturally it follows that $x \neq x_i$ for any $x_i \in E_\delta$. Additionally, we have that $d(x, x_i) \geq \delta$ for each $x_i \in E_\delta$. But then x should have been picked as an element of E_δ, a contradiction.

We now define the set

$$E = \bigcup_{n=1}^{\infty} E_{1/n}$$

Since each $E_{1/n}$ is finite (i.e., at most countable), the Corollary to Theorem 2.12 implies that E is at most countable. Additionally, since we may choose a different starting x_1 for each $E_{1/n}$, we can make E strictly countable. We now seek to prove that E is dense in X. Let $x \in X$ be arbitrary. If $x \in E$, we are done. If $x \notin E$, let $N_r(x)$ be an arbitrary neighborhood of x. By the Archimedean property, we may choose $1/n < r$ where $n \in \mathbb{N}$. We know from the above that $\{N_{1/n}(x_i)\}_{x_i\in E_{1/n}}$ is a cover of X, so there exists some $N_{1/n}(x_i)$ such that $x \in N_{1/n}(x_i)$. It follows that $x_i \in N_{1/n}(x) \subset N_r(x)$. This combined with the fact that $x_i \in E_{1/n} \subset E$ implies that x is a limit point of E, as desired. □

25. Prove that every compact metric space K has a countable base, and that K is therefore separable. (Hint: For every positive integer n, there are finitely many neighborhoods of radius $1/n$ whose union covers K.)

Proof. Let $n \in \mathbb{N}$ be arbitrary. Consider the collection $\{N_{1/n}(x)\}_{x\in K}$. Since each neighborhood is open and each $x \in K$ is in at least $N_{1/n}(x)$, $\{N_{1/n}(x)\}_{x\in K}$ is an open cover of K. It follows since K is compact that there exists a finite subcover $\{N_{1/n}(x_i)\}_{i=1}^{k_n}$.

We now define the set

$$V = \bigcup_{n=1}^{\infty} \{N_{1/n}(x_i)\}_{i=1}^{k_n}$$

Since each $\{N_{1/n}(x_i)\}_{i=1}^{k_n}$ is finite (i.e., at most countable), the Corollary to Theorem 2.12 implies that V is at most countable. Additionally, since the neighborhoods in each $\{N_{1/n}(x_i)\}_{i=1}^{k_n}$ have a distinct radius, we can pick an element in each $\{N_{1/n}(x_i)\}_{i=1}^{k_n}$ that is not in any other $\{N_{1/n}(x_i)\}_{i=1}^{k_n}$. This combined with the last result implies that K is strictly countable.

Moreover, we can show that V is a base of K. As previously mentioned, each $\{N_{1/n}(x_i)\}_{i=1}^{k_n}$ is a collection of open subsets of K, so V overall is also a collection of open subsets of K. Additionally, let $x \in K$ be arbitrary and let $G \subset K$ be an arbitrary open set containing x. Since G is open, there exists a neighborhood $N_r(x)$ such that $N_r(x) \subset G$. Using the Archimedean property, choose $1/n < r/2$. Since $\{N_{1/n}(x_i)\}_{i=1}^{k_n}$ is a cover of K, we know that $x \in N_{1/n}(x_i) \in V$ for some i. It follows that

$$x \in N_{1/n}(x_i) \subset N_r(x) \subset G$$

as desired.

Having established that K has a countable base, we now seek to prove that K is separable. Consider the set E containing the centers of each neighborhood in V. Since V is countable, E is countable. To complete the proof, we will show that every $x \in K$ is either an element of E or E', proving that E is dense in K. Let $x \in K$ be arbitrary. If $x \in E$, we are done. If $x \notin E$, let $N_r(x)$ be an arbitrary neighborhood of x. Use the Archimedean property to choose $1/n < r$. Since $\{N_{1/n}(x_i)\}_{i=1}^{k_n}$ covers K, $x \in N_{1/n}(x_i)$ for some i. Thus, $x_i \in N_{1/n}(x) \subset N_r(x)$ and $x_i \in E$ by definition, as desired. □

26. Let X be a metric space in which every infinite subset has a limit point. Prove that X is compact. (Hint: By Exercises 2.23 and 2.24, X has a countable base. It follows that every open cover of X has a *countable* subcover $\{G_n\}_{n\in\mathbb{N}}$. If no finite subcollection of $\{G_n\}$ covers X, then the complement F_n of $\bigcup_1^n G_i$ is nonempty for each n, but $\bigcap F_n$ is empty. If E is a set which contains a point from each F_n, consider a limit point of E, and obtain a contradiction.)

Proof. By Exercise 2.24, X is separable. It follows by Exercise 2.23 that X has a countable base.

To prove that X is compact, it will suffice to show that every open cover of X has a finite subcover. Consider an arbitrary open cover of X. Since X has a countable base, each element of this open cover is the union of an at most countable number of sets in the base. Thus, we may pick for each element of the base an element of the open cover of which it is a subset, and let this collection $\{G_n\}$ of elements of the open cover be a countable subcover of the open cover.

Now suppose for the sake of contradiction that $\{G_n\}$ is strictly countable and that no finite subset of it covers X. Then each $F_n = (\bigcup_1^n G_i)^c$ is nonempty (for otherwise the finite subset $\{G_1, \dots, G_n\}$ would cover X). Now let E be a set containing a distinct point from each F_n (since each F_n is nonempty, we can choose *a* point from each F_n, but moreover we can choose a *distinct* point from each F_n for otherwise, some F_n would be finite and we could choose a finite subcover of $\{G_n\}$ consisting of all sets picked up until a finite F_n is generated plus one set for each remaining point in F_n). As an infinite set, E has a limit point x by hypothesis. As an element of X, $x \in G_i$ for some $G_i \in \{G_n\}$. Thus, there exists $N_r(x) \subset G_i$. But since x is a limit point of E, there exists some point $e \in E$, not equal to x, in $N_r(x)$. Since $e \in G_i$, $e \notin F_j$ for $j \geq i$. Thus, $e \in F_j$ for some $j < i$. But since there are only finitely many F_j with $j < i$, there are only finitely many (at most $i-1$) $e \in N_r(x)$, contradicting Theorem 2.20's assertion that $N_r(x)$ contains infinitely many points of E. □

27. Define a point p in a metric space X to be a **condensation point** of a set $E \subset X$ if every neighborhood of p contains uncountably many points of E. Suppose $E \subset \mathbb{R}^k$ is uncountable, and let P be the set of all condensation points of E. Prove that P is perfect and that at most countably many points of E are not in P. In other words, show that $P^c \cap E$ is at most countable. (Hint: Let $\{V_n\}$ be a countable base of $\mathbb{R}^k$, let W be the union of those V_n for which $E \cap V_n$ is at most countable, and show that $P = W^c$.)

Proof. By Exercise 2.22, $\mathbb{R}^k$ is separable. Thus, by Exercise 2.23, $\mathbb{R}^k$ has a countable base $\{V_n\}$. Let W be the union of all V_n for which $E \cap V_n$ is at most countable. By the Corollary to Theorem 2.12, W is at most countable. Additionally, by Theorem 2.24, W is open.

We now seek to show that $P = W^c$. Let $x \in P$ be arbitrary. Suppose for the sake of contradiction that $x \in W$. Then $x \in V_n$ for some n where V_n is open. It follows that there exists $N_r(x) \subset V_n$. Since x is a condensation point of E, $N_r(x)$ contains uncountably many points of E. But this contradicts our hypothesis that $V_n \cap E$ is at most countable. Therefore, $x \in W^c$. Now suppose that $x \in W^c$. Let $N_r(x)$ be an arbitrary neighborhood of x. As an open set containing x, there exists some V_n in the countable base such that $x \in V_n \subset N_r(x)$. Since $x \in W^c$, x is not in any V_i for which $E \cap V_n = i$ is at most countable, i.e., $V_n \cap E$ must be uncountable. It follows since $V_n \subset N_r(x)$ that $N_r(x)$ contains uncountably many points of E, meaning that $x \in P$, as desired.

Having shown that $P = W^c$, we know that $W = P^c$. Thus, since W is at most countable, $P^c \cap E = W \cap E$ is at most countable, as desired. Lastly, having show that $P = W^c$ where W is open, we have by Theorem 2.23 that P is closed. Additionally, we can show that every $x \in P$ is a limit point of P. Let $x \in P$ be arbitrary, and let $N_r(x)$ be arbitrary. Since x is a condensation point of E, $N_r(x) \cap E$ is uncountable. It follows by the above that the set $\tilde{P}$ of all condensation points of $N_r(x) \cap E$ contains all but an at most countable number of points of $N_r(x) \cap E$, i.e., contains uncountably many points of E. Since $\tilde{P} \subset P$, there exist an element (indeed, uncountably many elements) of P in $N_r(x)$, as desired. Therefore, P is perfect, as desired. $\square$

28. Prove that every closed set in a separable metric space is the union of a (possibly empty) perfect set and a set which is at most countable. Corollary: Every countable closed set in $\mathbb{R}^k$ has isolated points. (Hint: Use Exercise 2.27.)

Proof. Let E be a closed subset of a separable metric space X. If E is at most countable, then $E = \emptyset \cup E$ is the union of an empty perfect set and an at most countable set. If E is uncountable, then since X is separable, Exercise 2.23 asserts that X has a countable base. Additionally, since E is closed, the set P of all condensation points of E is a subset of E. Thus, Exercise 2.27 implies that $E = P \cup (P^c \cap E)$ is the union of a perfect set and a set which is at most countable, as desired. $\square$

29. Prove that every open set in $\mathbb{R}^1$ is the union of an at most countable collection of disjoint segments (Hint: Use Exercise 22.)

Proof. Let E be an arbitrary open subset of $\mathbb{R}^1$. By Exercise 2.22, $\mathbb{R}^1$ is separable. Thus, it has a countable dense subset which we may call K. Consider $K \cap E$. Since K is countable, $K \cap E$ is at most countable. In particular, since E is open in $\mathbb{R}^1$, $K \cap E$ is strictly countable. Thus, let $K \cap E = \{x_1, x_2, \dots\}$. To $x_1 \in E$, assign the longest segment (a_1, b_1) such that $x_1 \in (a_1, b_1) \subset E$. Now if possible, choose x_i such that $x_i \notin (a_1, b_1)$ and $x_j \notin (a_1, b_1)$ implies $j \geq i$. To this x_i, assign the longest segment (a_i, b_i) such that $x_i \in (a_i, b_i) \subset E$.

Continuing on in this fashion forever will clearly yield a set of segments whose union is E (since K is dense in $\mathbb{R}^1$). From this set of segments there exists an obvious injection into K countable, so this set of segments is at most countable. Additionally, by the construction, every $(a_i, b_i) \cap (a_j, b_j) = \emptyset$ for WLOG $i < j$ (for if otherwise, then we would have chosen the segment $(\min(a_i, a_j), \max(b_i, b_j))$ when dealing with x_i). $\square$

30. Imitate the proof of Theorem 2.43 to obtain the following result: If $\mathbb{R}^k = \bigcup_1^\infty F_n$, where each F_n is a closed subset of $\mathbb{R}^k$, then at least one F_n has a nonempty interior. Equivalent statement: If G_n is a dense open subset of $\mathbb{R}^k$ for each $n \in \mathbb{N}$, then $\bigcap_1^\infty G_n$ is not empty (in fact, it is dense in $\mathbb{R}^k$). This is a special case of Baire's theorem; see Exercise 3.22 for the general case.

Proof. We will prove the first statement herein. Let's begin.

Suppose for the sake of contradiction that $F_n^\circ = \emptyset$ for all $n \in \mathbb{N}$. Then

$$\begin{aligned} \emptyset &= \bigcup_{i=1}^\infty F_n^\circ \\ \mathbb{R}^k &= \left(\bigcup_{i=1}^\infty F_n^\circ\right)^c \\ &= \bigcap_{i=1}^\infty (F_n^\circ)^c && \text{Theorem 2.22} \\ &= \bigcap_{i=1}^\infty \overline{F_n^c} && \text{Exercise 2.9d} \end{aligned}$$

It follows that $F_n^c = \mathbb{R}^k$ for all $n \in \mathbb{N}$. This combined with the fact that each F_n^c is open (since each F_n is closed by hypothesis) implies that each F_n^c is dense in $\mathbb{R}^k$. Now let G be a nonempty open subset of $\mathbb{R}^k$. Since F_1^c is dense in $\mathbb{R}^k$, there exists $p \in F_1^c$ such that $p \in G$ (G nonempty means there exists $x \in G$; G open means there exists $N_r(x) \subset G$; F_1^c dense implies there exists $p \in F_1^c$ such that $p \in N_r(x) \subset G$). Let $G_1 = F_1^c \cap G$. $G_1 \neq \emptyset$ by the previous result. Additionally, since F_1^c and G are both open, Theorem 2.24c implies that G_1 is open. Thus, we may choose $p_1 \in G_1$ and know that there exists a $N_{2r_1}(p_1) \subset G_1$. It follows that $\overline{N_{r_1}(p_1)} \subset G_1$. In much the same way, we can construct $G_2 = F_2^c \cap N_{r_1}(p_1)$ and find $p_2 \in G_2$ such that there exists $\overline{N_{r_2}(p_2)} \subset G_2$. Continuing in this fashion yields a decreasing sequence of compact sets $\overline{N_{r_1}(p_1)} \supset \overline{N_{r_2}(p_2)} \supset \cdots$. Thus, by Theorem 2.39, $\bigcap_{n=1}^\infty \overline{N_{r_n}(p_n)} \neq \emptyset$. Thus, since each $\overline{N_{r_n}(p_n)} \subset F_n^c$, $\bigcap_{n=1}^\infty \overline{N_{r_n}(p_n)} \subset \bigcap_{n=1}^\infty F_n^c$, meaning that

$$\begin{aligned} \emptyset &\neq \bigcap_{n=1}^\infty F_n^c \\ &= \left(\bigcup_{n=1}^\infty F_n\right)^c && \text{Theorem 2.22} \\ \mathbb{R}^k &\neq \bigcup_{n=1}^\infty F_n \end{aligned}$$

a contradiction. $\square$

8 Continuity

From Rudin (1976).

Chapter 4

11/29: **1.** Suppose f is a real function defined on $\mathbb{R}^1$ which satisfies

$$\lim_{h\to 0}[f(x+h)-f(x-h)] = 0$$

for every $x \in \mathbb{R}^1$. Does this imply that f is continuous?

Proof. No.

Consider the function $f : \mathbb{R}^1 \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} 1 & x = 0 \\ 0 & x \neq 0 \end{cases}$$

We first show that f satisfies the desired property. We divide into two cases ($x = 0$ and $x \neq 0$). If $x = 0$, then since $f(0+h) = 0$ for all $h \neq 0$ by the definition of f, we have that for every $\epsilon > 0$, there exists δ (arbitrarily choose $\delta = 1$) such that $|[f(x+h)-f(x-h)]-0| = |[0-0]-0| = 0 < \epsilon$ for all $h \in \mathbb{R}^1$ satisfying $0 < |h-0| < \delta$. It follows that $\lim_{h\to 0}[f(0+h)-f(0-h)] = 0$, as desired. On the other hand, if $x \neq 0$, assume WLOG that $x > 0$ (the argument is symmetric if $x < 0$). Choose δ such that $0 < \delta < x$. Then $\delta < |x-0|$, so $x \pm h \neq 0$ for any h satisfying $0 < |h| < \delta$ (for otherwise, we would have $h = \pm x$ and thus $|x| < \delta$, contradicting $\delta < |x|$). It follows that for every $\epsilon > 0$, there exists δ (this chosen δ) such that $|[f(x+h)-f(x-h)]-0| = 0 < \epsilon$ for all $h \in \mathbb{R}^1$ satisfying $0 < |h-0| < \delta$.

Second, we show that f is not continuous. Specifically, f is not continuous at 0 since the fact that $f(y) = 0$ for any $y \neq 0$ implies that $\lim_{y\to 0} f(y) = 0 \neq 1 = f(x)$. ☐

2. If f is a continuous mapping of a metric space X into a metric space Y, prove that

$$f(\bar{E}) \subset \overline{f(E)}$$

for every set $E \subset X$ ($\bar{E}$ denotes the closure of E). Show, by an example, that $f(\bar{E})$ can be a proper subset of $\overline{f(E)}$.

Proof. Let $f(x) \in f(\bar{E})$ be arbitrary. It follows by the definition of images that $x \in \bar{E}$. We now divide into two cases ($x \in E$ and $x \in E'$). If $x \in E$, then $f(x) \in f(E) \subset f(E) \cup f(E)' = \overline{f(E)}$ as desired. On the other hand, if $x \in E'$, then every neighborhood of x contains some element of E other than x. We now look to show that every neighborhood of $f(x)$ contains some element of $f(E)$ other than $f(x)$. Let $N_\epsilon(f(x))$ be an arbitrary neighborhood of $f(x)$. Since $f(x)$ is continuous, we have that $\lim_{y\to x} f(y) = f(x)$. It follows that there exists $\delta > 0$ such that $d_X(y,x) < \delta$ implies $d_Y(f(y), f(x)) < \epsilon$. By the hypothesis that $x \in E'$, $N_\delta(x)$ contains some $y \in E$ such that $y \neq x$. Moreover, since $y \in E$, $f(y) \in f(E)$. Additionally, the previous statement implies that $d_Y(f(y), f(x)) < \epsilon$. If $f(x) = f(y)$, then $f(x) \in f(E) \subset \overline{f(E)}$, and if $f(x) \neq f(y)$, then our neighborhood $N_\epsilon(f(x))$ contains an element of $f(E)$ other than $f(x)$, as desired.

Let $f : (0,1) \to \mathbb{R}$ be defined by $f(x) = x$. Then

$$f[\overline{(0,1)}] = f[(0,1)] = (0,1) \subsetneq [0,1] = \overline{f[(0,1)]}$$

as desired. ☐

3. Let f be a continuous real function on a metric space X. Let $Z(f)$ (the **zero set** of f) be the set of all $p \in X$ at which $f(p) = 0$. Prove that $Z(f)$ is closed.

Proof. By definition, $Z(f) = f^{-1}(\{0\})$. Thus, since $\{0\}$ is closed as a finite set and f is continuous, the Corollary to Theorem 4.8 implies that $Z(f)$ is closed. $\square$

4. Let f and g be continuous mappings of a metric space X into a metric space Y, and let E be a dense subset of X. Prove that $f(E)$ is dense in $f(X)$. If $g(p) = f(p)$ for all $p \in E$, prove that $g(p) = f(p)$ for all $p \in X$. (In other words, a continuous mapping is determined by its values on a dense subset of its domain.)

Proof. To prove that $f(E)$ is dense in $f(X)$, it will suffice to show that every $f(x) \in f(X)$ is either an element or a limit point of $f(E)$. Let $f(x) \in f(X)$ be arbitrary. If $f(x) \in f(E)$, we are done. If $f(x) \notin f(E)$, then $x \notin E$. It follows however by the density of E in X that x is a limit point of E. Therefore,

$$\begin{aligned} f(x) &\in f(\bar{E}) \\ &\subset \overline{f(E)} && \text{Exercise 4.2} \\ &= f(E) \cup f(E)' \end{aligned}$$

so since $f(x) \notin f(E)$, we must have $f(x) \in f(E)'$, as desired.

As to the other part of the proof, suppose for the sake of contradiction that there exists $p \in X$ such that $g(p) \neq f(p)$. Now since $g(p) \neq f(p)$, $d_Y(g(p), f(p)) \neq 0$. In particular, we may let $d_Y(g(p), f(p)) = 2\epsilon$ where $\epsilon > 0$. It follows by the continuity of g that there exists $\delta_1 > 0$ such that $d_X(x, p) < \delta_1$ implies $d_Y(g(x), g(p)) < \epsilon$, and symmetrically by the continuity of f that there exists $\delta_2 > 0$ such that $d_X(x, p) < \delta_2$ implies $d_Y(f(x), f(p)) < \epsilon$. Choose $\delta = \min(\delta_1, \delta_2)$. Since E is dense in X, there exists $x \in E$ such that $x \in N_\delta(p)$. Consequently, since $d_X(x, p) < \delta \leq \delta_1$, we have that $d_Y(g(x), g(p)) < \epsilon$, and symmetrically that $d_Y(f(x), f(p)) < \epsilon$. But since $x \in E$, $f(x) = g(x)$. Therefore, we have that

$$d_Y(g(p), f(p)) \leq d_Y(g(p), g(x)) + d_Y(f(x), f(p)) < \epsilon + \epsilon = 2\epsilon$$

contradicting the previously proven fact that $d_Y(g(p), f(p)) = 2\epsilon$. $\square$

5. If f is a real continuous function defined on a closed set $E \subset \mathbb{R}^1$, prove that there exist continuous real functions g on $\mathbb{R}^1$ such that $g(x) = f(x)$ for all $x \in E$. (Such functions g are called **continuous extensions** of f from E to $\mathbb{R}^1$.) Show that the result becomes false if the word "closed" is omitted. Extend the result to vector-valued functions. (Hint: Let the graph of g be a straight line on each of the segments which constitute the complement of E [compare Exercise 2.29]. The result remains true if $\mathbb{R}^1$ is replaced by any metric space, but the proof is not so simple.)

Proof. Since E is closed, E^c is open. Thus, by Exercise 2.29, E^c is the union of an at most countable collection of disjoint segments. In particular, we may let

$$E^c = \bigcup_{i=1}^{n} (a_i, b_i)$$

where $n \in [1, \infty]$ and $(a_i, b_i) \cap (a_j, b_j) = \emptyset$ for any $i \neq j$. Thus, we may let $g : \mathbb{R}^1 \to \mathbb{R}$ be defined by

$$g(x) = \begin{cases} f(x) & x \in E \\ f(a_i) + \frac{f(b_i) - f(a_i)}{b_i - a_i}(x - a_i) & x \in (a_i, b_i) \subset E^c \end{cases}$$

Clearly $g(x) = f(x)$ for all $x \in E$. All that remains is to show that g is continuous. To do so, we divide into three cases ($x \in E^\circ$, $x \in E^c$, and $x \in E \setminus E^\circ$).

First, suppose $x \in E^\circ$. To prove that g is continuous at x, it will suffice to show that for every $\epsilon > 0$, there exists a $\delta > 0$ such that $|g(y) - g(x)| < \epsilon$ for all $y \in \mathbb{R}^1$ for which $|y - x| < \delta$. Since $x \in E^\circ$, there exists $N_{\delta_1}(x) \subset E$. Additionally, since f is continuous at x, there exists a $\delta_2 > 0$ such that

$|f(y) - f(x)| < \epsilon$ for all $y \in E$ for which $|y - x| < \delta_2$. Choose $\delta = \min(\delta_1, \delta_2)$. Now suppose $y \in \mathbb{R}^1$ and $|y - x| < \delta$. Since $|y - x| < \delta \leq \delta_1$, $y \in N_{\delta_1}(x) \subset E$. This combined with the fact that $x \in E$ by hypothesis implies that $g(y) = f(y)$ and $g(x) = f(x)$. It follows since $|y - x| < \delta \leq \delta_2$ that

$$\begin{aligned}|g(y) - g(x)| &= |f(y) - f(x)| \\ &< \epsilon\end{aligned}$$

as desired.

Second, suppose $x \in E^c$. Then $x \in (a_i, b_i)$ for some $i \in \mathbb{N}$. To prove that g is continuous at x, it will suffice to show that for every $\epsilon > 0$, there exists a $\delta > 0$ such that $|g(y) - g(x)| < \epsilon$ for all $y \in \mathbb{R}^1$ for which $|y - x| < \delta$. Let $\epsilon > 0$ be arbitrary. Since $x \in (a_i, b_i)$ open, there exists $N_{\delta_1}(x) \subset (a_i, b_i)$. Additionally, let $\delta_2 = \epsilon \cdot |(b_i - a_i)/(f(b_i) - f(a_i))|$. Let $\delta = \min(\delta_1, \delta_2)$. Now suppose $y \in \mathbb{R}^1$ and $|y - x| < \delta$. Since $|y - x| < \delta \leq \delta_1$, $y \in N_{\delta_1}(x) \subset (a_i, b_i)$. This combined with the fact that $x \in (a_i, b_i)$ by hypothesis implies that

$$g(y) = f(a_i) + \frac{f(b_i) - f(a_i)}{b_i - a_i}(y - a_i) \qquad g(x) = f(a_i) + \frac{f(b_i) - f(a_i)}{b_i - a_i}(x - a_i)$$

It follows since $|y - x| < \delta \leq \delta_2$

$$\begin{aligned}|g(y) - g(x)| &= \left|\left[f(a_i) + \frac{f(b_i) - f(a_i)}{b_i - a_i}(y - a_i)\right] - \left[f(a_i) + \frac{f(b_i) - f(a_i)}{b_i - a_i}(x - a_i)\right]\right| \\ &= \left|\frac{f(b_i) - f(a_i)}{b_i - a_i}(y - x)\right| \\ &= \left|\frac{f(b_i) - f(a_i)}{b_i - a_i}\right| \cdot |y - x| \\ &< \left|\frac{f(b_i) - f(a_i)}{b_i - a_i}\right| \cdot \epsilon \cdot \left|\frac{b_i - a_i}{f(b_i) - f(a_i)}\right| \\ &= \epsilon\end{aligned}$$

as desired.

Third, suppose $x \in E \setminus E^\circ$. We now show that this means that $x = a_i$ or $x = b_j$ for some i, j, or that $x \in E'$. If $x = a_i$ or $x = b_j$ for some i, j, then clearly $x \notin E^c$ (for otherwise there would be a segment (a_k, b_k) containing it that is not disjoint from (a_i, b_i) [resp. (a_j, b_j)]), i.e., $x \in E$. Additionally, $x \notin E^\circ$ since it's status as the endpoint of a segment means that there are points of E^c arbitrarily close to it. On the other hand, if $x \in E \setminus E^\circ$ and $x \neq a_i, b_j$ for any i, j, then we can show that $x \in E'$. Indeed, consider $N_r(x)$. Since $x \notin E^\circ$, there exists a point $y \in E^c$ such that $y \in N_r(x)$. Suppose for the sake of definiteness that $y > x$ (the proof is symmetric if $y < x$). It follows since $y \in E^c$ and $x \notin E^c$ that there exits a segment (a_k, b_k) containing y but not containing x. Naturally, this must imply that $x < a_k < y < b_k$. But since each $a_k \in E$ as previously established, and $|a_k - x| < |y - x| < r$, $a_k \in N_r(x)$, as desired. Having established that $x = a_i$ or $x = b_j$ for some i, j, or that $x \in E'$, we now divide into these two subcases. In particular, for the first subcase, we divide into three subsubcases ($x = a_i$ and $x \neq b_j$, $x \neq a_i$ and $x = b_j$, and $x = a_i = b_j$).

Suppose first that $x = a_i$ for some i and $x \neq b_j$ for any j. To prove that g is continuous at x, it will suffice to show that $g(x+) = g(x-) = g(x)$. Since $x = a_i$, we can show by an argument analogous to that used in the second case that $g(x+) = f(a_i) = g(x)$. Additionally, since $x \neq b_j$ for any j, there exists $(y, x] \subset E$ for some $y < x$. Thus, we can show by an argument analogous to that used in the first case that $g(x-) = f(x-) = f(x) = g(x)$.

The proof of the second subsubcase is symmetric to that of the first.

For the third subsubcase, simply apply the first part of the proof of the first subsubcase twice, once to each "side" of x.

For the second subcase, we know that if $x \in E \setminus E^\circ$, then $x \in E$, so $\lim_{y \to x} f(y) = f(x)$. This combined with the fact that $x \in E'$ implies that y can actually approach x. Moreover, we know by the definition of g that if f is bounded in sufficiently small regions of x (as it is), g will be bounded in sufficiently small regions of x by the same bounds (if points of g exceeded the minimal bounds of f, then there would have to be some points of f not extended/connected to each other via a straight line). Therefore, g is continuous at x, as desired. ☐

6. Let (X, d_X) and (Y, d_Y) be metric spaces, and let $f : E \to Y$ where E is a compact subset of X. Consider the **graph** $G \subset X \times Y$ of f, where the metric on $X \times Y$ is $d = d_X + d_Y$, i.e., $d[(x_1, y_1), (x_2, y_2)] = d_X(x_1, x_2) + d_Y(y_1, y_2)$. Show that f is continuous if and only if G is compact. (Hint: There are several ways of doing this. There is a "topological" proof that only uses the fact that compact sets are closed in a metric space, and the fact that a function is continuous if and only if pre-images of closed sets are closed. Another way to go about it is to use sequential compactness [i.e., any sequence contained in a compact set has a convergent subsequence].)

Proof. Suppose first that f is continuous. Define $\mathbf{g} : E \to X \times Y$ by $\mathbf{g}(x) = (x, f(x))$ for all $x \in E$. Then since the component functions g_1, g_2 are continuous, Theorem 4.10 implies that $\mathbf{g}$ is continuous. Therefore, since $\mathbf{g}$ is continuous and E is compact, Theorem 4.14 implies that $\mathbf{g}(E) = G$ is compact.

Now suppose that G is compact. To prove that f is continuous, the Corollary to Theorem 4.8 tells us that it will suffice to show that for every closed set C in Y, $f^{-1}(C)$ is closed in X. Let C be an arbitrary closed set in Y. Define $\pi_1 : G \to X$ and $\pi_2 : G \to Y$ by

$$\pi_1(x, f(x)) = x \qquad\qquad \pi_2(x, f(x)) = f(x)$$

We can prove that π_1, π_2 are continuous by choosing $\delta = \epsilon$ in the definition of continuity. It follows by the Corollary to Theorem 4.8 that $\pi_2^{-1}(C)$ is closed in G. This combined with the fact that G is compact implies that $\pi_2^{-1}(C)$ is compact. This combined with the fact that π_1 is continuous implies by Theorem 4.14 that $\pi_1(\pi_2^{-1}(C)) = f^{-1}(C)$ is compact. Thus, by Theorem 2.34, $f^{-1}(C)$ is closed, as desired. ☐

7. If $E \subset X$ and if f is a function defined on X, the **restriction** of f to E is the function g whose domain of definition is E such that $g(p) = f(p)$ for $p \in E$. Define f and g on $\mathbb{R}^2$ by

$$f(x,y) = \begin{cases} 0 & (x,y) = (0,0) \\ \frac{xy^2}{x^2+y^4} & (x,y) \neq (0,0) \end{cases} \qquad\qquad g(x,y) = \begin{cases} 0 & (x,y) = (0,0) \\ \frac{xy^2}{x^2+y^6} & (x,y) \neq (0,0) \end{cases}$$

Prove that f is bounded on $\mathbb{R}^2$, that g is unbounded in every neighborhood of $(0,0)$, and that f is not continuous at $(0,0)$; nevertheless, the restrictions of both f and g to every straight line in $\mathbb{R}^2$ are continuous!

Proof. <u>f is bounded on $\mathbb{R}^2$</u>: To prove that f is bounded on $\mathbb{R}^2$, it will suffice to show that there exists a real number M such that $|f(x,y)| \leq M$ for all $(x,y) \in \mathbb{R}^2$. Choose $M = 1/2$. Let $(x,y) \in \mathbb{R}^2$ be arbitrary. Then

$$\begin{aligned} 0 &\leq (y^2 - x)^2 \\ 0 &\leq y^4 - 2xy^2 + x^2 \\ 2xy^2 &\leq x^2 + y^4 \\ \frac{xy^2}{x^2+y^4} &\leq \frac{1}{2} \end{aligned}$$

as desired.

<u>g is unbounded in every neighborhood of $(0,0)$</u>: Let $N_r(0,0)$ be an arbitrary neighborhood of $(0,0)$. Suppose for the sake of contradiction that there exists a real number M such that $|g(x,y)| \leq M$ for

all $(x,y) \in N_r(0,0)$. Let $n_1 \in \mathbb{N}$ be such that $10^{n_1} > 2M$. Let $n_2 \in \mathbb{N}$ be such that $(10^{-3n_2}, 10^{-n_2}) \in N_r(0,0)$. Let $n = \max(n_1, n_2)$; note that this implies that $10^n > 2M$ and $(10^{-3n}, 10^{-n}) \in N_r(0,0)$. Then by the latter statement and the fact that

$$\begin{aligned} M &< \frac{1}{2} \cdot 10^n \\ &= \frac{10^{-5n}}{2 \cdot 10^{-6n}} \\ &= \frac{(10^{-3n})(10^{-n})^2}{(10^{-3n})^2 + (10^{-n})^6} \\ &= g(10^{-3n}, 10^{-n}) \\ &\leq |g(10^{-3n}, 10^{-n})| \end{aligned}$$

we have found an $(x,y) \in N_r(0,0)$ such that $|g(x,y)| > M$, a contradiction.

<u>f is not continuous at $(0,0)$</u>: To prove that f is not continuous at $(0,0)$, it will suffice to show that there exists an $\epsilon > 0$ such that for all $\delta > 0$, there exists $(x,y) \in \mathbb{R}^2$ such that $\|(x,y) - (0,0)\| < \delta$ and $|f(x,y) - f(0,0)| \geq \epsilon$. Choose $\epsilon = 1/2$. Let $\delta > 0$ be arbitrary. Choose $(y^2, y) \in \mathbb{R}^2$ such that $\left\|(y^2,y) - (0,0)\right\| = \left\|(y^2,y)\right\| < \delta$. This combined with the fact that

$$\begin{aligned} |f(y^2,y) - f(0,0)| &= \left| \frac{(y^2)(y)^2}{(y^2)^2 + (y)^4} \right| \\ &= \left| \frac{y^4}{2y^4} \right| \\ &= \left| \frac{1}{2} \right| \\ &= \epsilon \end{aligned}$$

completes the proof.

<u>The restriction of f to any straight line in $\mathbb{R}^2$ is continuous</u>: We divide into two cases (straight lines of the form $y = ax + b$ where $a, b \in \mathbb{R}$, and straight lines of the form $x = c$ where $c \in \mathbb{R}$). In the first case, let $\tilde{f} : \{(x,y) : y = ax + b\} \to \mathbb{R}$ be the restriction of f to the arbitrary straight line $y = ax + b$ of the first form. Then

$$\tilde{f}(x,y) = \tilde{f}(x, ax+b) = \frac{x(ax+b)^2}{x^2 + (ax+b)^4}$$

for every $(x,y) \in \{(x,y) : y = ax + b\}$. Since the rightmost function above is the result of sums, products, and quotients of the continuous functions $x \mapsto x$ and $x \mapsto ax + b$, Theorem 4.9 asserts that $\tilde{f}$ is continuous on its domain, except possibly when $x^2 + (ax+b)^4 = 0$. However, this will only happen in the special case when $b = 0$ and $(x,y) = (x, ax+0) = (x, ax) = (0,0)$. Thus, to complete the proof, we need only show that for every $\epsilon > 0$, there exists a $\delta > 0$ such that $|\tilde{f}(x,ax) - \tilde{f}(0,0)| < \epsilon$ if $\|(x,ax) - (0,0)\| < \delta$. Let $\epsilon > 0$ be arbitrary. Choose $\delta = \epsilon/a^2$ if $a \neq 0$ (and $\delta = 1$ for the trivial case where $a = 0$ and thus $\tilde{f} = 0$ as well). Let (x, ax) such that $\|(x,ax) - (0,0)\| = \|(x,ax)\| < \delta$ be arbitrary. This importantly implies that $|x| < \delta$. Therefore,

$$\begin{aligned} |\tilde{f}(x,ax) - \tilde{f}(0,0)| &= \left| \frac{(x)(ax)^2}{(x)^2 + (ax)^4} \right| \\ &= \left| \frac{a^2x^3}{a^4x^4 + x^2} \right| \\ &= \left| \frac{a^2x}{a^4x^2 + 1} \right| \\ &\leq \left| \frac{a^2x}{1} \right| \end{aligned}$$

$$\begin{aligned} &= a^2 \cdot |x| \\ &< a^2 \cdot \frac{\epsilon}{a^2} \\ &= \epsilon \end{aligned}$$

as desired.

In the second case, let $\tilde{f} : \{(c, y)\} \to \mathbb{R}$ be the restriction of f to an arbitrary straight line $x = c$ of the second form. A symmetric argument to the other case completes the proof.

The restriction of g to any straight line in $\mathbb{R}^2$ is continuous: The proof is symmetric to the above. $\square$

8. Let f be a real uniformly continuous function on the bounded set E in $\mathbb{R}^1$. Prove that f is bounded on E. Show that the conclusion is false if boundedness of E is omitted from the hypothesis.

Proof. Since E is dense in $\bar{E}$ and f is a uniformly continuous real function on E, Exercise 4.13 asserts that f has a continuous extension g from E to $\bar{E}$. Since $E \subset \mathbb{R}^1$ is bounded, $\bar{E} \subset \mathbb{R}^1$ is closed and bounded, and hence compact by the Heine-Borel theorem. This combined with the fact that g is continuous implies by Theorem 4.14 that $g(\bar{E})$ is compact. Thus, $g(\bar{E})$ must be closed and bounded by the Heine-Borel theorem, so $f(E) \subset g(\bar{E})$ must be bounded. It follows trivially that f is bounded.

Let $E = \mathbb{R}^1$ and $f : E \to \mathbb{R}$ be defined by $f(x) = x$ for all $x \in E$. Then f is a real uniformly continuous function on E unbounded for which $f(E) = E$ is naturally unbounded. $\square$

9. Show that the requirement in the definition of uniform continuity can be rephrased as follows, in terms of diameters of sets: To every $\epsilon > 0$, there exists a $\delta > 0$ such that $\operatorname{diam} f(E) < \epsilon$ for all $E \subset X$ with $\operatorname{diam} E < \delta$.

Proof. Suppose first that $f : X \to Y$ is uniformly continuous, where $(X, d_X), (Y, d_Y)$ are metric spaces. We wish to prove that to every $\epsilon > 0$, there exists a $\delta > 0$ such that $\operatorname{diam} f(E) < \epsilon$ for all $E \subset X$ with $\operatorname{diam} E < \delta$. Let $\epsilon > 0$ be arbitrary. Since f is uniformly continuous, Definition 4.18 tells us that there exists $\delta > 0$ such that $d_Y(f(p), f(q)) < \epsilon/2$ for all $p, q \in X$ for which $d_X(p, q) < \delta$. Choose this δ to be our δ. Let E be an arbitrary subset of X satisfying $\operatorname{diam} E < \delta$. To prove that $\operatorname{diam} f(E) < \epsilon$, it will suffice to show that

$$\sup\{d_Y(f(p), f(q)) : p, q \in E\} < \epsilon$$

Since $\operatorname{diam} E < \delta$, $d_X(p, q) < \delta$ for all $p, q \in E$. Thus, for all $p, q \in E \subset X$, $d_Y(f(p), f(q)) < \epsilon/2$. It follows that

$$\operatorname{diam} f(E) = \sup\{d_Y(f(p), f(q)) : p, q \in E\} \leq \frac{\epsilon}{2} < \epsilon$$

as desired.

The proof is symmetric in the other direction. $\square$

10. Complete the details of the following alternative proof of Theorem 4.19: If f is not uniformly continuous, then for some $\epsilon > 0$, there are sequences $\{p_n\}, \{q_n\}$ in X such that $d_X(p_n, q_n) \to 0$ but $d_Y(f(p_n), f(q_n)) > \epsilon$. Use Theorem 2.37 to obtain a contradiction.

Proof. Let $(X, d_X), (Y, d_Y)$ where the former is compact, and let $f : X \to Y$ be continuous. Now suppose for the sake of contradiction that f is not uniformly continuous. Then there exists a number $2\epsilon > 0$ such that for all $\delta > 0$, there exist $p, q \in X$ for which $d_X(p, q) < \delta$ but $d_Y(f(p), f(q)) \geq 2\epsilon$. Let $\{\delta_n\}_1^\infty$ be defined by $\delta_n = 1/n$. For each δ_n, use the above statement to choose $p_n, q_n \in X$ satisfying $d_X(p_n, q_n) < \delta_n$ and $d_Y(f(p_n), f(q_n)) \geq 2\epsilon$. It follows that $\{p_n\}, \{q_n\}$ are sequences in X such that $d_X(p_n, q_n) \to 0$ but $d_Y(f(p_n), f(q_n)) > \epsilon$. Moreover, the ranges of $\{p_n\}, \{q_n\}$ are infinite. (Suppose otherwise. Then there would be a pair of terms p_n, q_n that are the closest together. Let these terms be separated by a distance r. We know that $r \neq 0$ since $f(p_n), f(q_n)$ are separated by a nonzero distance, hence are distinct, and f, as a function, cannot map the same input to distinct outputs. Moreover, there would then not be a pair of terms corresponding to $\delta_m = 1/m < r$, which we know to exist by the Archimedean principle, a contradiction.)

Since f is continuous and X is compact, Theorem 4.14 implies that $f(X)$ is compact. Thus, since the ranges of $\{p_n\}, \{q_n\}, \{f(p_n)\}, \{f(q_n)\}$ are infinite subsets of compact sets, they have limit points. Hence, each of these four sequences converges. Moreover, $\{p_n\}, \{q_n\}$ converge to the same point $p = q$ since $d_X(p_n, q_n) \to 0$ while $\{f(p_n)\}, \{f(q_n)\}$ converge to different points since $d_Y(f(p_n), f(q_n)) > \epsilon$. However, this contradicts the continuity of f at $p = q$ since it proves the existence of points arbitrarily close by that map to separate elements of Y. □

11. Suppose f is a uniformly continuous mapping of a metric space X into a metric space Y and prove that $\{f(x_n)\}$ is a Cauchy sequence in Y for every Cauchy sequence $\{x_n\}$ in X. Use this result to give an alternative proof of the theorem stated in Exercise 4.13.

Proof. Let $\{x_n\}$ be an arbitrary Cauchy sequence in X. To prove that $\{f(x_n)\}$ is a Cauchy sequence in Y, it will suffice to show that for every $\epsilon > 0$, there exists an integer N such that $d_Y(f(x_n), f(x_m)) < \epsilon$ if $n, m \geq N$. Let $\epsilon > 0$ be arbitrary. Since f is uniformly continuous, there exists a $\delta > 0$ such that $d_Y(f(p), f(q)) < \epsilon$ for all $p, q \in X$ satisfying $d_X(p, q) < \delta$. Moreover, since $\{x_n\}$ is Cauchy in X, there exists an integer N such that $d_X(x_n, x_m) < \delta$ if $n, m \geq N$. Choose this N to be our N. Let $n, m \geq N$ be arbitrary. Then since $d_X(x_n, x_m) < \delta$, it follows by the above that $d_Y(f(x_n), f(x_m)) < \epsilon$, as desired.

Suppose $E \subset X$ is dense in (X, d_X) and let $f : E \to \mathbb{R}$ be uniformly continuous. We wish to prove that f has a continuous extension $g : X \to \mathbb{R}$.

We first define an extension g of f as follows. For $x \in E$, let $g(x) = f(x)$. For $x \notin E$, choose a sequence $\{x_n\}$ in E that converges to x (we know that one exists by the density of E in X). Since $\{x_n\}$ is convergent, Theorem 3.11a implies that it's Cauchy. Since $\{x_n\}$ is Cauchy, we have by the above that $\{f(x_n)\}$ is Cauchy. It follows by Theorem 3.11c that $\{f(x_n)\}$ converges to a point in $\mathbb{R}$ that we may define to be $g(x)$.

We now prove that g as defined is continuous. Suppose for the sake of contradiction that g is not continuous at some $x \in X$. Then there exists an $\epsilon > 0$ such that for all $\delta > 0$, there exists $y \in X$ satisfying $d(x, y) < \delta$ and $d_Y(g(x), g(y)) \geq \epsilon$. We use this statement to define a sequence $\{y_n\}$ of points in X, none of which is equal to x, that converges to x. First, let $\{\delta_n\}_1^\infty$ be defined by $\delta_n = 1/n$. Then let $\{y_n\}_1^\infty$ be a sequence where each y_n satisfies $y_n \in E$, $d(x, y_n) < \delta_n$, and $d_Y(g(x), g(y_n)) \geq \epsilon$; we know such a point exists for each n by the above condition and by the density of E in X, and that none of the points equals x since there is a nonzero distance between $g(x)$ and $g(y_n)$ and g is a function (i.e., cannot have multiple definitions on one object). Since $\delta_n \to 0$, $y_n \to x$. However, $g(y_n) \nrightarrow g(x)$. If $x \in E$, then this means that we can find points of E arbitrarily close to x that nevertheless map to values isolated from $g(x) = f(x)$, contradicting the continuity of f. If $x \notin E$, then $g(x)$ is *defined* to be the limit of $\{y_n\}$, leading to a contradiction with the definition of g. □

12. A uniformly continuous function of a uniformly continuous function is uniformly continuous. State this more precisely and prove it.

Proof. Let $(X, d_X), (Y, d_Y), (Z, d_Z)$ be metric spaces, and let $E \subset X$. Suppose $f : E \to Y$ and $g : f(E) \to Z$ are uniformly continuous. Then the composition $h : E \to Z$ of f and g is uniformly continuous.

To prove that h is uniformly continuous, it will suffice to show that for every $\epsilon > 0$, there exists $\delta > 0$ such that $d_Z(h(x), h(x')) < \epsilon$ for all $x, x' \in E$ satisfying $d_X(x, x') < \delta$. Let $\epsilon > 0$ be arbitrary. Since g is uniformly continuous, there exists $\eta > 0$ such that $d_Z(g(y), g(y')) < \epsilon$ for all $y, y' \in f(E)$ satisfying $d_Y(y, y') < \eta$. Since f is uniformly continuous, there exists $\delta > 0$ such that $d_Y(f(x), f(x')) < \eta$ for all $x, x' \in E$ satisfying $d_X(x, x') < \delta$. Let x, x' be arbitrary points of E that satisfy $d_X(x, x') < \delta$. Then $d_Y(f(x), f(x')) < \eta$. It follows that

$$d_Z(h(x), h(x')) = d_Z(g(f(x)), g(f(x'))) < \epsilon$$

as desired. □

13. Let E be a dense subset of a metric space X and let f be a uniformly continuous *real* function defined on E. Prove that f has a continuous extension from E to X (see Exercise 4.5 for terminology). Uniqueness follows from Exercise 4.4. (Hint: For each $p \in X$ and each positive integer n, let $V_n(p)$ be the set of all $q \in E$ with $d(p,q) < 1/n$. Use Exercise 4.9 to show that the intersection of the closures of the sets $f(V_1(p)), f(V_2(p)), \dots$ consists of a single point, say $g(p)$, of $\mathbb{R}^1$. Prove that the function g so defined on X is the desired extension of f.) Could the range space $\mathbb{R}^1$ be replaced by $\mathbb{R}^k$? By any compact metric space? By any complete metric space? By any metric space?

Proof. We begin by defining a function g on X. After defining it, we will prove it is a continuous extension of f from E to X. Let's begin.

Let $p \in X$ be arbitrary. Define the family of sets $\{V_n(p)\}_1^\infty$ by $V_n(p) = \{q \in E : d(q,p) < 1/n\}$ for all $n \in \mathbb{N}$. We now show that the intersection of the images of every set in this collection under f contains exactly one point in $\mathbb{R}$ that we may define to be $g(p)$. First off, note that by definition, $V_n(p) \supset V_{n+1}(p)$ for all $n \in \mathbb{N}$. Thus, $f(V_n(p)) \supset f(V_{n+1}(p))$ and hence $\overline{f(v_n(p))} \supset \overline{f(V_{n+1}(p))}$ for all $n \in \mathbb{N}$. It follows that $\bigcap_{n\in\mathbb{N}} \overline{f(V_n(p))} = \bigcap_{k\in K} \overline{f(V_k(p))}$ for any infinite $K \subset \mathbb{N}$. Indeed, to prove our desired result that $\bigcap_{n\in\mathbb{N}} \overline{f(V_n(p))}$ is a singleton set, we need only show that the intersection of some infinite subset of $\{\overline{f(V_n(p))}\}$ is a singleton set. We may do so via Theorem 3.10b; the invocation of said result requires that we construct an infinite collection $\{\overline{f(V_{k_n}(p))}\} \subset \{\overline{f(V_n(p))}\}$ of compact, decreasing sets with $\lim_{n\to\infty} \operatorname{diam} \overline{f(V_{k_n}(p))} = 0$. We will perform the construction first, and then confirm that it meets the three criteria.

Let $\epsilon_n = 1/n$ for each $n \in \mathbb{N}$. Since f is uniformly continuous, Exercise 4.9 tells us that there exists a $\delta_n > 0$ such that $\operatorname{diam} f(F) < \epsilon_n$ for all $F \subset E$ with $\operatorname{diam} F < \delta_n$ for each $n \in \mathbb{N}$. By consecutive applications of the Archimedean principle and using strong induction, to each δ_n, assign a $k_n \in \mathbb{N}$ such that $2/k_n < \delta_i$ $(i = 1, \dots, n)$ and $k_n > k_i$ $(i = 1, \dots, n-1)$. Let $K = \{k_n : n \in \mathbb{N}\}$. This completes the construction. Now for the check, let $n \in \mathbb{N}$ be arbitrary.

To confirm that $\overline{f(V_{k_n}(p))}$ is compact, the Heine-Borel theorem tells us that it will suffice to demonstrate that $\overline{f(V_{k_n}(p))}$ is closed and bounded. By Theorem 2.27a, $\overline{f(V_{k_n}(p))}$ is closed, as desired. Since $\operatorname{diam} V_{k_n}(p) = 2/k_n < \delta_n$, we have by the above that $\operatorname{diam} f(V_{k_n}(p)) < \epsilon_n$. Hence $\operatorname{diam} \overline{f(V_{k_n}(p))} \leq \epsilon_n$, verifying that $\overline{f(V_{k_n}(p))}$ is bounded, as desired.

To confirm that $\overline{f(V_{k_n}(p))} \supset \overline{f(V_{k_{n+1}}(p))}$, it will suffice to show that $V_{k_n}(p) \supset V_{k_{n+1}}(p)$. But we know this to be true by the definition of $V_n(p)$ and the fact that $k_n < k_{n+1}$ by the construction, as desired.

To confirm that $\lim_{n\to\infty} \operatorname{diam} \overline{f(V_{k_n}(p))} = 0$, we will use the squeeze theorem. In particular, since $\lim_{n\to\infty} 0 = 0$, $\lim_{n\to\infty} \epsilon_n = 0$, and $0 \leq \operatorname{diam} \overline{f(V_{k_n}(p))} \leq \epsilon_n$ for all $n \in \mathbb{N}$, we must have that $\lim_{n\to\infty} \operatorname{diam} \overline{f(V_{k_n}(p))} = 0$, as desired.

Having established that

$$\left| \bigcap_{n\in\mathbb{N}} \overline{f(V_n(p))} \right| = \left| \bigcap_{k_n\in K} \overline{f(V_{k_n}(p))} \right| = 1$$

we may define $\{g(p)\} = \bigcap_{n\in\mathbb{N}} \overline{f(V_n(p))}$.

We now seek to prove that $g(p) = f(p)$ for all $p \in E$. Suppose for the sake of contradiction that for some $p \in E$, $g(p) \neq f(p)$. Let $|g(p) - f(p)| = 2\epsilon > 0$. Since f is uniformly continuous, Exercise 4.9 asserts that there exists a $\delta > 0$ such that $\operatorname{diam} f(F) < \epsilon$ for all $F \subset E$ with $\operatorname{diam} F < \delta$. Now use the Archimedean principle to choose $2/m < \delta$. It follows by the definition of diameter that $\operatorname{diam} V_m(p) < \delta$. Thus, by the above condition, $\operatorname{diam} \overline{f(V_m(p))} < \epsilon$. Additionally, since $g(p) \in \bigcap_{n\in\mathbb{N}} \overline{f(V_n(p))}$, $g(p) \in \overline{f(V_m(p))}$, and since $p \in V_n(p)$, $f(p) \in \overline{f(V_m(p))}$. But this implies by the definition of diameter that $|g(p) - f(p)| \leq \epsilon < 2\epsilon$, a contradiction.

A symmetric argument to the above proves that g is continuous (specifically, if there is a discontinuity at $g(p)$, then we can find an $\overline{f(V_m(p))}$ of which $g(p)$ is not an element, contradicting the way $g(p)$ is defined).

As to the other questions, the only place where we make use of the properties of $\mathbb{R}$ is when we use the Heine-Borel theorem to assert that any closed and bounded set is compact. We can still make this logical step in $\mathbb{R}^k$ (Theorem 2.41), in compact metric spaces (Theorem 2.35), and in complete metric spaces (Definition 3.12), but not in general metric spaces (Exercise 2.16). □

14. Let $I = [0, 1]$ be the closed unit interval. Suppose f is a continuous mapping of I into I. Prove that $f(x) = x$ for at least one $x \in I$.

Proof. Consider the function $g : I \to \mathbb{R}$ defined by $g(x) = f(x) - x$. To prove that $f(x) = x$ for some $x \in I$, it will suffice to show that $g(x) = 0$ for some $x \in I$. We divide into two cases ($g(0) = 0$ or $g(1) = 0$, and $g(0) \neq 0$ and $g(1) \neq 0$). In the first case, we are done immediately. In the second case, we have

$$\begin{aligned} g(0) = f(0) - 0 &\neq 0 & \qquad g(1) = f(1) - 1 &\neq 0 \\ f(0) &\neq 0 & f(1) &\neq 1 \end{aligned}$$

It follows since $f(0), f(1) \in [0, 1]$ that $f(0) > 0$ and $f(1) < 1$. Thus, $g(0) = f(0) - 0 > 0$ and $g(1) = f(1) - 1 < 0$. Additionally, since f and $x \mapsto x$ are continuous, Theorem 4.9 asserts that the difference of them (i.e., g) is continuous as well. Therefore, since g is a continuous real function on the interval $[0, 1]$ and $g(0) > 0 > g(1)$, Theorem 4.23 asserts that there exists a point $x \in (0, 1)$ such that $g(x) = 0$, as desired. □

15. Call a mapping of X into Y open if $f(V)$ is an open set in Y whenever V is an open set in X. Prove that every continuous open mapping of $\mathbb{R}^1$ into $\mathbb{R}^1$ is monotonic.

Proof. We will first prove that f is 1-1. Suppose $x \neq y$, and WLOG let $x < y$. We seek to demonstrate that $f((x, y)) = (c, d)$ where $c, d \in \mathbb{R}$ are distinct, and that $f([x, y]) = [c, d]$; it will follow that $f(x) = c$ and $f(y) = d$ or vice versa, proving either way that $f(x) \neq f(y)$ as desired. Let's begin. To demonstrate the first claim, it will suffice to show that $f((x, y))$ is open, connected, and bounded. Since f is open, $f((x, y))$ is open. Since f is continuous and (x, y) is connected, Theorem 4.22 asserts that $f((x, y))$ is connected. Since f is continuous and $[x, y]$ is compact, Theorem 4.19 asserts that f is uniformly continuous on $[x, y]$; hence f is bounded on (x, y) by Exercise 4.8, as desired. On the other hand, to demonstrate the second claim, it will suffice to show that $f([x, y])$ is compact and connected. Since f is continuous and $[x, y]$ is compact, Theorem 4.14 implies that $f([x, y])$ is compact. Since f is continuous and $[x, y]$ is connected, Theorem 4.22 implies that $f([x, y])$ is connected, as desired.

Now suppose for the sake of contradiction that f is not monotonic. Then there exist $x < y < z$ such that $f(x) < f(y)$ and $f(y) > f(z)$ (or, symmetrically, such that $f(x) > f(y)$ and $f(y) < f(z)$). Since f is 1-1 and $x \neq z$, $f(x) \neq f(z)$. We divide into two cases ($f(x) < f(z)$ and $f(x) > f(z)$). If $f(x) < f(z)$, then $f(x) < f(z) < f(y)$ by hypothesis. Thus, since f is a continuous real function on $[x, y]$ and $f(x) < f(z) < f(y)$, Theorem 4.23 asserts that there exists a $c \in (x, y)$ such that $f(c) = f(z)$. But since f is 1-1, this implies that $c = z$, meaning that $x < z < y$, a contradiction. The proof of the other case is symmetric. □

16. Let $[x]$ denote the largest integer contained in x, that is, $[x]$ is the integer such that $x - 1 < [x] \leq x$; and let $(x) = x - [x]$ denote the fractional part of x. What discontinuities do the functions $[x]$ and (x) have?

Proof. $[x]$ has a simple discontinuity at each $z \in \mathbb{Z}$. At each $z \in \mathbb{Z}$, $f(z+) = z$ and $f(z-) = z - 1$. At each $x \notin \mathbb{Z}$, $f(x+) = f(x-) = [x]$.

(x) also has a simple discontinuity at each $z \in \mathbb{Z}$. At each $z \in \mathbb{Z}$, $f(z+) = 0$ and $f(z-) = 1$. At each $x \notin \mathbb{Z}$, $f(x+) = f(x-) = (x)$. □

17. Let f be a real function defined on (a, b). Prove that the set of points at which f has a simple discontinuity is at most countable. (Hint: Let E be the set on which $f(x-) < f(x+)$. With each point x of E, associate a triple (p, q, r) of rational numbers such that

(a) $f(x-) < p < f(x+)$;

(b) $a < q < t < x$ implies $f(t) < p$;

(c) $x < t < r < b$ implies $f(t) > p$.

The set of all such triples is countable. Show that each triple is associated with at most one point of E. Deal similarly with the other possible types of simple discontinuities.)

Proof. Let D be the set of points at which f has a simple discontinuity. Let E be the set of all $x \in D$ such that $f(x-) < f(x+)$. Let $x \in E$ be arbitrary. We now show that we can choose a (p, q, r) pertaining to x as described in the hint. By the density of $\mathbb{Q} \subset \mathbb{R}$, choose $p \in \mathbb{Q}$ such that $f(x-) < p < f(x+)$. Let $\epsilon = p - f(x-)$. Since $\lim_{t \to x^-} f(t) = f(x-)$, there exists a $\delta > 0$ such that if $t \in (a, b)$ and $0 < x - t < \delta$, then $|f(t) - f(x-)| < \epsilon = p - f(x-)$; in particular, $f(t) < p$. Choose $q \in \mathbb{Q}$ such that $x - \delta < q < x$. Choose r symmetrically, as per the hint.

We now show that if $y \in E$ such that (p, q, r) pertain to y, then $y = x$. Suppose for the sake of contradiction that $y \neq x$. WLOG let $x < y$. Choose $q < x < t < y < r$. It follows since $q < t < y$ that $f(t) < p$. It follows since $x < t < r$ that $f(t) > p$, a contradiction.

We can treat the set F of all $x \in D$ such that $f(x-) > f(x+)$ symmetrically.

Having defined an injective function from both E and F to $\mathbb{Q}^3$, we know that E and F are at most countable. Thus, their union (D) is also at most countable, as desired. ☐

18. Every rational x can be written in the form $x = m/n$, where $n > 0$ and m and n are integers without any common divisors. When $x = 0$, we take $n = 1$. Consider the function f defined on $\mathbb{R}^1$ by

$$f(x) = \begin{cases} 0 & x \notin \mathbb{Q} \\ \frac{1}{n} & x = \frac{m}{n} \end{cases}$$

Prove that f is continuous at every irrational point, and that f has a simple discontinuity at every rational point.

19. Suppose f is a real function with domain $\mathbb{R}^1$ which has the intermediate value property: If $f(a) < c < f(b)$, then $f(x) = c$ for some x between a and b. Suppose also, for every rational r, that the set of all x with $f(x) = r$ is closed. Prove that f is continuous. (Hint: If $x_n \to x_0$, but $f(x_n) > r > f(x_0)$ for some r and all n, then $f(t_n) = r$ for some t_n between x_0 and x_n; thus, $t_n \to x_0$. Find a contradiction. (Fine, 1966).)

20. If E is a nonempty subset of a metric space X, define the **distance** from $x \in X$ to E by

$$\rho_E(x) = \inf_{z \in E} d(x, z)$$

(a) Prove that $\rho_E(x) = 0$ if and only if $x \in \bar{E}$.

Proof. Suppose first that $\rho_E(x) = 0$. Let $N_r(x)$ be an arbitrary neighborhood of x. Since $\inf_{z \in E} d(x, z) = 0$, there exists a $z \in E$ such that $0 \leq d(x, z) < r$. This $z \in E$ will therefore be an element of $N_r(x)$, proving that $x \in \bar{E}$, as desired. Now suppose that $x \in \bar{E}$. Then there is some point of E in every neighborhood of x. Thus, since there are elements $z \in E$ arbitrarily close to x, there are elements $z \in E$ that make $d(x, z)$ arbitrarily small. Thus, $\rho_E(x) = \inf_{z \in E} d(x, z) = 0$, as desired. ☐

(b) Prove that ρ_E is a uniformly continuous function on X, by showing that

$$|\rho_E(x) - \rho_E(y)| \leq d(x, y)$$

for all $x, y \in X$. (Hint: $\rho_E(x) \leq d(x, z) \leq d(x, y) + d(y, z)$, so that $\rho_E(x) \leq d(x, y) + \rho_E(y)$.)

Proof. Let $x, y \in X$ be arbitrary. We have that

$$\rho_E(x) \leq d(x,z) \leq d(x,y) + d(y,z)$$

for all $z \in E$ by the definition of ρ_E. In particular, considering a sequence $\{z_n\}$ in E such that $d(y, z_n) \to \inf_{z\in E} d(y,z)$ yields

$$\begin{aligned} \rho_E(x) &\leq d(x,y) + \rho_E(y) \\ \rho_E(x) - \rho_E(y) &\leq d(x,y) \end{aligned}$$

Interchanging the roles of x and y in the above algebra yields

$$\rho_E(y) - \rho_E(x) \leq d(y,x) = d(x,y)$$

Therefore, we have that

$$|\rho_E(x) - \rho_E(y)| \leq d(x,y)$$

It follows that if we want $|\rho_E(x) - \rho_E(y)| < \epsilon$, we need only require that $d(x,y) < \delta = \epsilon$, so ρ_E is uniformly continuous, as desired. ☐

21. Suppose K compact and F closed are disjoint sets in a metric space X. Prove that there exists $\delta > 0$ such that $d(p,q) > \delta$ if $p \in K$, $q \in F$. (Hint: ρ_F is a continuous positive function on K.) Show that the conclusion may fail for two disjoint closed sets if neither is compact.

Proof. Since F is closed, $F = \bar{F}$. Thus, by Exercise 4.20a, $\rho_F(x) = 0$ if and only if $x \in F$. It follows since K is disjoint from F that $\rho_F(x) \neq 0$ for all $x \in K$. In particular, since the distance function is strictly nonnegative, ρ must be strictly nonnegative, meaning that $\rho_F(x) > 0$ for all $x \in K$. Additionally, we have by Exercise 4.20b that ρ_F is uniformly continuous on X. Thus, since ρ_F is continuous and K is compact, Theorem 4.14 asserts that $\rho_F(K)$ is compact. This combined with the previous result implies that $0 \notin \rho_F(K)$ and, since $\rho_F(K)$ is closed as a compact set, 0 is isolated from $\rho_F(K)$. Consequently, there exists $2\delta > 0$ such that $N_{2\delta}(0) \cap \rho_F(K) = \emptyset$. It follows that $\rho_F(p) \geq 2\delta > \delta$ for all $p \in K$. Therefore, if $p \in K$ and $q \in F$, then

$$d(p,q) \geq \rho_F(p) > \delta$$

as desired.

As a counterexample, consider the sets A, B of all rational numbers less than $\sqrt{2}$ and all rational numbers greater than $\sqrt{2}$, respectively. A and B are both closed and disjoint, but since we can find rational numbers arbitrarily close to $\sqrt{2}$ from both sides, the minimum distance between two points in the sets converges to zero. ☐

22. Let A and B be disjoint nonempty closed sets in a metric space X, and define

$$f(p) = \frac{\rho_A(p)}{\rho_A(p) + \rho_B(p)}$$

for all $p \in X$. Show that f is a continuous function on X whose range lies in $[0,1]$, that $f(p) = 0$ precisely on A, and that $f(p) = 1$ precisely on B. This establishes a converse of Exercise 4.3: Every closed set $A \subset X$ is $Z(f)$ for some continuous real f on X. Setting

$$V = f^{-1}([0, \tfrac{1}{2})) \qquad W = f^{-1}((\tfrac{1}{2}, 1])$$

show that V and W are open and disjoint, and that $A \subset V$, $B \subset W$. (Thus pairs of disjoint closed sets in a metric space can be covered by pairs of disjoint open sets. This property of metric spaces is called **normality**.)

Proof. Since f is the result of sums and quotients of continuous (Exercise 4.20b) functions, Theorem 4.9 asserts that f is continuous, except possibly where $\rho_A(p)+\rho_B(p)=0$. However, this will never be the case: Since both functions in the sum are nonnegative, the sum can only be zero if both functions are equal to zero. But if $\rho_A(p)=0$ and $\rho_B(p)=0$, then $p\in A$ and $\rho_B(p)=0$ by Exercise 4.20a, contradicting the fact that A,B are disjoint. Thus, f is everywhere continuous, as desired.

Since ρ_A,ρ_B are nonnegative, $0\leq\rho_A(p)\leq\rho_A(p)+\rho_B(p)$ for all $p\in X$. Thus,

$$0\leq\frac{\rho_A(p)}{\rho_A(p)+\rho_B(p)}=f(p)\leq 1$$

for all $p\in X$, as desired.

If $p\in A$, then

$$\begin{aligned}
f(p)&=\frac{\rho_A(p)}{\rho_A(p)+\rho_B(p)}\\
&=\frac{0}{0+\rho_B(p)} && \text{Exercise 4.20a}\\
&=0
\end{aligned}$$

as desired. On the other hand, if $f(p)=0$, then $\rho_A(p)=0$, so Exercise 4.20a implies that $p\in A$, as desired.

If $p\in B$, then

$$\begin{aligned}
f(p)&=\frac{\rho_A(p)}{\rho_A(p)+\rho_B(p)}\\
&=\frac{\rho_A(p)}{\rho_A(p)+0} && \text{Exercise 4.20a}\\
&=1
\end{aligned}$$

as desired. On the other hand, if $f(p)=1$, then $\rho_B(p)=0$, so Exercise 4.20a implies that $p\in B$, as desired.

Since $[0,\frac{1}{2})$ is open in $[0,1]$ and f is continuous, Theorem 4.8 asserts that $V=f^{-1}([0,\frac{1}{2}))$ is open in X. Similarly, W is open in X. Additionally, V,W are disjoint since if $x\in V\cap W$, then $f(x)<1/2$ and $f(x)>1/2$, a contradiction. Lastly, if $p\in A$, the $f(p)=0$, so $p\in V$. Similarly, $B\subset W$. $\square$

23. A real-valued function f defined in (a,b) is said to be **convex** if

$$f(\lambda x+(1-\lambda)y)\leq\lambda f(x)+(1-\lambda)f(y)$$

whenever $a<x<b$, $a<y<b$, and $0<\lambda<1$. Prove that every convex function is continuous. Prove that every increasing convex function of a convex function is convex. (For example, if f is convex, so is e^f.) If f is convex in (a,b) and if $a<s<t<u<b$, show that

$$\frac{f(t)-f(s)}{t-s}\leq\frac{f(u)-f(s)}{u-s}\leq\frac{f(u)-f(t)}{u-t}$$

24. Assume that f is a continuous real function defined in (a,b) such that

$$f\left(\frac{x+y}{2}\right)\leq\frac{f(x)+f(y)}{2}$$

for all $x,y\in(a,b)$. Prove that f is convex.

Proof. Suppose for the sake of contradiction that f is not convex. Then there exist $c,d\in(a,b)$ and $\lambda\in(0,1)$ such that $f(\lambda c+(1-\lambda)d)>\lambda f(c)+(1-\lambda)f(d)$. WLOG let $c<d$ (we know $c\neq d$ since the two sides of the convexity condition would be equal were c equal to d). Consider $g:(a,b)\to\mathbb{R}$ defined by

$$g(x)=f(x)-\left[f(c)+\frac{f(d)-f(c)}{d-c}(x-c)\right]$$

Analogously to f, we have that

$$\begin{aligned}
g\left(\frac{x+y}{2}\right)&=f\left(\frac{x+y}{2}\right)-\left[f(c)+\frac{f(d)-f(c)}{d-c}\left(\frac{x+y}{2}-c\right)\right]\\
&=f\left(\frac{x+y}{2}\right)-\left[\frac{f(c)+f(c)}{2}+\frac{f(d)-f(c)}{d-c}\left(\frac{x-c+y-c}{2}\right)\right]\\
&=f\left(\frac{x+y}{2}\right)-\frac{1}{2}\left[\left(f(c)+\frac{f(d)-f(c)}{d-c}(x-c)\right)+\left(f(c)+\frac{f(d)-f(c)}{d-c}(y-c)\right)\right]\\
&\leq\frac{f(x)+f(y)}{2}-\frac{1}{2}\left[\left(f(c)+\frac{f(d)-f(c)}{d-c}(x-c)\right)+\left(f(c)+\frac{f(d)-f(c)}{d-c}(y-c)\right)\right]\\
&=\frac{1}{2}\left[\left(f(x)-\left(f(c)+\frac{f(d)-f(c)}{d-c}(x-c)\right)\right)+\left(f(y)-\left(f(c)+\frac{f(d)-f(c)}{d-c}(y-c)\right)\right)\right]\\
&=\frac{g(x)+g(y)}{2}
\end{aligned}$$

for all $x,y\in(a,b)$, and that g is continuous as the result of sums and products of continuous functions (Theorem 4.9). Additionally, we know that $g(c)=g(d)=0$ (by definition) and that

$$\begin{aligned}
g(\lambda c+(1-\lambda)d)&=f(\lambda c+(1-\lambda)d)-\left[f(c)+\frac{f(d)-f(c)}{d-c}([\lambda c+(1-\lambda)d]-c)\right]\\
&=f(\lambda c+(1-\lambda)d)-\left[f(c)+\frac{f(d)-f(c)}{d-c}((1-\lambda)d-(1-\lambda)c)\right]\\
&=f(\lambda c+(1-\lambda)d)-[f(c)+(1-\lambda)(f(d)-f(c))]\\
&=f(\lambda c+(1-\lambda)d)-f(c)-(1-\lambda)f(d)+(1-\lambda)f(c)\\
&>\lambda f(c)+(1-\lambda)f(d)-f(c)-(1-\lambda)f(d)+(1-\lambda)f(c)\\
&=0
\end{aligned}$$

Since g is continuous and $[c,d]$ is compact, Theorem 4.16 asserts that g attains its maximum, say of $g(e)$ at $e\in[c,d]$. It follows since $g(\lambda c+(1-\lambda)d)>0$ that $g(e)\geq g(\lambda c+(1-\lambda)d)>0$. We now divide into two cases ($d(c,e)\leq d(e,d)$ and $d(c,e)>d(e,d)$). In the first case, let $\delta=d(c,e)$. Since $g(e)\geq g(x)$ for all $x\in[c,d]$, we know that $g(c+2\delta)\leq g(e)$. It follows since $g(e)>0$ that

$$g(e)=g(c+\delta)=g\left(\frac{c+(c+2\delta)}{2}\right)\leq\frac{g(c)+g(c+2\delta)}{2}=\frac{g(c+2\delta)}{2}<g(c+2\delta)\leq g(e)$$

a contradiction. The proof is symmetric in the other case. ☐

25. If $A,B\subset\mathbb{R}^k$, define $A+B$ to be the set of all sums $\mathbf{x}+\mathbf{y}$ with $\mathbf{x}\in A$, $\mathbf{y}\in B$.

(a) If K is compact and C is closed in $\mathbb{R}^k$, prove that $K+C$ is closed. (Hint: Take $\mathbf{z}\notin K+C$, put $F=\mathbf{z}-C$, the set of all $\mathbf{z}-\mathbf{y}$ with $\mathbf{y}\in C$. Then K and F are disjoint. Choose δ as in Exercise 4.21. Show that the open ball with center $\mathbf{z}$ and radius δ does not intersect $K+C$.)

Proof. Suppose first that $K+C=\mathbb{R}^k$. Then $K+C$ is closed.
Having dealt with the trivial case, we now seek to prove that $K+C\subsetneq\mathbb{R}^k$ is closed. We will do so by proving that $(K+C)^c$ is open. To do so, it will suffice to show that to every $\mathbf{z}\in(K+C)^c$ there corresponds a $N_\delta(\mathbf{z})\subset(K+C)^c$. Let $\mathbf{z}\in(K+C)^c$ be arbitrary. Define $F=\mathbf{z}-C$.

Now suppose for the sake of contradiction that $\mathbf{a} \in K \cap F$. Then $\mathbf{a} = \mathbf{z} - \mathbf{b}$ for some $\mathbf{b} \in C$. Additionally, since $\mathbf{z} \notin K + C$, $\mathbf{z} \neq \mathbf{x} + \mathbf{y}$ for any $\mathbf{x} \in K$, $\mathbf{y} \in C$. Thus, $\mathbf{a} \neq (\mathbf{x} + \mathbf{b}) - \mathbf{b} = \mathbf{x}$ for any $\mathbf{x} \in K$, so $\mathbf{a} \notin K$, a contradiction. Therefore, K, F are disjoint. It follows since K compact and F closed are disjoint subsets of the metric space $\mathbb{R}^k$ by Exercise 4.21 that there exists $\delta > 0$ such that $\|\mathbf{x} - \mathbf{y}\| > \delta$ for all $\mathbf{x} \in K$, $\mathbf{y} \in F$. Now suppose there exists $\mathbf{x} + \mathbf{y} \in N_\delta(\mathbf{z})$ where $\mathbf{x} \in K$ and $\mathbf{y} \in C$. Then $\|\mathbf{x} + \mathbf{y} - \mathbf{z}\| = \|\mathbf{x} - (\mathbf{z} - \mathbf{y})\| < \delta$, a contradiction. ☐

(b) Let α be an irrational real number. Let C_1 be the set of all integers, and let C_2 be the set of all $n\alpha$ with $n \in C_1$. Show that C_1 and C_2 are closed subsets of $\mathbb{R}^1$ whose sum $C_1 + C_2$ is *not* closed, by showing that $C_1 + C_2$ is a countable dense subset of $\mathbb{R}^1$.

Proof. C_1 and C_2 are both closed since every point in each is isolated.

$C_1 + C_2$ is countable since we can define a natural injection from it to $C_1 \times C_2$ and we know that cross products of countable sets are countable. $C_1 + C_2$ is dense in $\mathbb{R}^1$ since the set of all fractional parts of all $n\alpha \in C_2$ is dense in $[0, 1]$ because α is irrational (and thus there is no repeating cycle), and we can shift this dense segment using values in C_1. Thus, since $C_1 + C_2$ is a countable dense subset of $\mathbb{R}^1$, $\overline{C_1 + C_2} = \mathbb{R}^1$ is uncountable, i.e., contains more elements than $C_1 + C_2$, showing that $C_1 + C_2$ is not closed. ☐

26. Suppose X, Y, Z are metric spaces, and Y is compact. Let $f : X \to Y$, let $g : Y \to Z$ be continuous and 1-1, and let $h(x) = g(f(x))$ for all $x \in X$. Prove that f is uniformly continuous if h is uniformly continuous. (Hint: g^{-1} has compact domain $g(Y)$, and $f(x) = g^{-1}(h(x))$.) Prove also that f is continuous if h is continuous. Show (by modifying Example 4.21, or by finding a different example) that the compactness of Y cannot be omitted from the hypotheses, even when X and Z are compact.

Proof. To prove that f is uniformly continuous, it will suffice to show that for every $\epsilon > 0$, there exists a $\delta > 0$ such that if $x, y \in X$ and $d_X(x, y) < \delta$, then $d_Y(f(x), f(y)) < \epsilon$. Let $\epsilon > 0$ be arbitrary. Since g is continuous and 1-1 on Y compact, Theorem 4.17 implies that $g^{-1} : g(Y) \to Y$ is continuous. Additionally, since g is continuous and Y is compact, $g(Y)$ is compact. The last two results imply by Theorem 4.19 that g^{-1} is uniformly continuous. Thus, there exists an $\eta > 0$ such that if $z, t \in g(Y)$ and $d_Z(z, t) < \eta$, then $d_Y(g^{-1}(z), g^{-1}(t)) < \epsilon$. Additionally, since h is uniformly continuous, there exists a $\delta > 0$ such that if $x, y \in X$ and $d_X(x, y) < \delta$, then $d_Z(h(x), h(y)) < \eta$. The previous two results combined with the fact that $h(x), h(y) \in g(Y)$ imply that

$$d_Y(f(x), f(y)) = d_Y(g^{-1}(h(x)), g^{-1}(h(y))) < \epsilon$$

as desired.

The proof that f is continuous given that h is continuous is symmetric to the above, except that it must be done pointwise for an each $x \in X$.

Let $X = [0, 2\pi]$, $Y = [0, 2\pi)$, and $Z = \{(r, \theta) \in \mathbb{R}^2 : r = 1, \theta \in \mathbb{R}\}$ be metric spaces under the normal Euclidean metric. Clearly X and Z are compact while Y is not. Let $f : X \to Y$ be defined by

$$f(x) = \begin{cases} \pi - x & x \in [0, \pi] \\ 3\pi - x & x \in (\pi, 2\pi] \end{cases}$$

Let $g : Y \to Z$ be defined by

$$g(x) = (\cos x, \sin x)$$

By Example 4.21, g is continuous and 1-1. Clearly $h = g \circ f$ is uniformly continuous while f is not. ☐

9 Sequences and Series of Functions

From Rudin (1976).

Chapter 7

12/10: **1.** Prove that every uniformly convergent sequence of bounded functions is uniformly bounded.

3. Construct sequences $\{f_n\}, \{g_n\}$ which converge uniformly on some set E, but such that $\{f_n g_n\}$ does not converge uniformly on E (of course, $\{f_n g_n\}$ must converge on E).

4. Consider

$$f(x) = \sum_{n=1}^{\infty} \frac{1}{1+n^2 x}$$

For what values of x does the series converge absolutely? On what intervals does it converge uniformly? On what intervals does it fail to converge uniformly? Is f continuous wherever the series converges? Is f bounded?

5. Let

$$f_n(x) = \begin{cases} 0 & x < \frac{1}{n+1} \\ \sin^2 \frac{\pi}{x} & \frac{1}{n+1} \leq x \leq \frac{1}{n} \\ 0 & \frac{1}{n} < x \end{cases}$$

Show that $\{f_n\}$ converges to a continuous function, but not uniformly. Use the series $\sum f_n$ to show that absolute convergence, even for all x, does not imply uniform convergence.

6. Prove that the series

$$\sum_{n=1}^{\infty} (-1)^n \frac{x^2+n}{n^2}$$

converges uniformly in every bounded interval, but does not converge absolutely for any value of x.

7. For $n = 1, 2, 3, \dots$ and x real, put

$$f_n(x) = \frac{x}{1+nx^2}$$

Show that $\{f_n\}$ converges uniformly to a function f and that the equation

$$f'(x) = \lim_{n\to\infty} f_n'(x)$$

is correct if $x \neq 0$ but false if $x = 0$.

8. If

$$I(x) = \begin{cases} 0 & x \leq 0 \\ 1 & x > 0 \end{cases}$$

if $\{x_n\}$ is a sequence of distinct points of (a, b), and if $\sum |c_n|$ converges, prove that the series

$$f(x) = \sum_{n=1}^{\infty} c_n I(x - x_n)$$

converges uniformly on $[a, b]$, and that f is continuous for every $x \neq x_n$.

9. Let $\{f_n\}$ be a sequence of continuous functions which converges uniformly to a function f on a set E. Prove that

$$\lim_{n\to\infty} f_n(x_n) = f(x)$$

for every sequence of points $x_n \in E$ such that $x_n \to x$ and $x \in E$. Is the converse of this true?

10. Letting (x) denote the fractional part of the real number x (see Exercise 4.16 for the definition), consider the function

$$f(x) = \sum_{n=1}^{\infty} \frac{(nx)}{n^2}$$

defined on $\mathbb{R}$. Find all discontinuities of f, and show that they form a countable dense set. Show that f is nevertheless Riemann-integrable on every bounded interval.

11. Suppose $\{f_n\}, \{g_n\}$ are defined on E and that

(a) $\sum f_n$ has uniformly bounded partial sums;

(b) $g_n \to 0$ uniformly on E;

(c) $g_1(x) \geq g_2(x) \geq g_3(x) \geq \cdots$ for every $x \in E$.

Prove that $\sum f_n g_n$ converges uniformly on E. (Hint: Compare with Theorem 3.42.)

12. Suppose g and f_n $(n \in \mathbb{N})$ are defined on $(0, \infty)$, are Riemann-integrable on $[t, T]$ whenever $0 < t < T < \infty$, $|f_n| \leq g$, $f_n \to f$ uniformly on every compact subset of $(0, \infty)$, and

$$\int_0^\infty g(x)\,\mathrm{d}x < \infty$$

Prove that

$$\lim_{n \to \infty} \int_0^\infty f_n(x)\,\mathrm{d}x = \int_0^\infty f(x)\,\mathrm{d}x$$

(See Exercises 6.7-6.8 for the relevant definitions.) This is a rather weak form of Lebesgue's dominated convergence theorem (Theorem 11.32). Even in the context of the Riemann integral, uniform convergence can be replaced by pointwise convergence if it is assumed that $f \in \mathscr{R}$. (See Cunningham (1967) and Kestelman (1970).)

13. Assume that $\{f_n\}$ is a sequence of monotonically increasing functions on $\mathbb{R}^1$ with $0 \leq f_n(x) \leq 1$ for all x and all n.

(a) Prove that there is a function f and a sequence $\{n_k\}$ such that

$$f(x) = \lim_{k \to \infty} f_{n_k}(x)$$

for every $x \in \mathbb{R}^1$. The existence of such a pointwise convergent subsequence is usually called **Helly's selection theorem**. (Hint: (i) Some subsequence $\{f_{n_i}\}$ converges at all rational points r, say, to $f(r)$. (ii) Define $f(x)$ for any $x \in \mathbb{R}^1$ to be $\sup f(r)$, the sup being taken over all $r \leq x$. (iii) Show that $f_{n_i}(x) \to f(x)$ at every x at which f is continuous. [This is where monotonicity is strongly used.] (iv) A subsequence of $\{f_{n_i}\}$ converges at every point of discontinuity of f since there are at most countably many such points.)

(b) If, moreover, f is continuous, prove that $f_{n_k} \to f$ uniformly on compact sets. (Hint: Modify your proof of (iii) appropriately.)

14. Let f be a continuous real function on $\mathbb{R}^1$ with the following properties: $0 \leq f(t) \leq 1$, $f(t+2) = f(t)$ for every t, and

$$f(t) = \begin{cases} 0 & 0 \leq t \leq \frac{1}{3} \\ 1 & \frac{2}{3} \leq t \leq 1 \end{cases}$$

Put $\Phi(t) = (x(t), y(t))$, where

$$x(t) = \sum_{n=1}^{\infty} 2^{-n} f(3^{2n-1} t) \qquad\qquad y(t) = \sum_{n=1}^{\infty} 2^{-n} f(3^{2n} t)$$

Prove that Φ is continuous and that Φ maps $I = [0,1]$ onto the unit square $I^2 \subset \mathbb{R}^2$. In fact, show that Φ maps the Cantor set onto I^2. (Hint: Each $(x_0, y_0) \in I^2$ has the form

$$x_0 = \sum_{n=1}^{\infty} 2^{-n} a_{2n-1} \qquad\qquad y_0 = \sum_{n=1}^{\infty} 2^{-n} a_{2n}$$

where each a_i is 0 or 1. If

$$t_0 = \sum_{i=1}^{\infty} 3^{-i-1}(2a_i)$$

show that $f(3^k t_0) = a_k$, and hence that $x(t_0) = x_0$, $y_0(t_0) = y_0$.) This simple example of a so-called **space-filling curve** is due to Schoenberg (1938).

15. Suppose f is a real continuous function on $\mathbb{R}^1$, $f_n(t) = f(nt)$ for $n \in \mathbb{N}$, and $\{f_n\}$ is equicontinuous on $[0,1]$. What conclusion can you draw about f?

16. Suppose $\{f_n\}$ is an equicontinuous sequence of functions on a compact set K, and $\{f_n\}$ converges pointwise on K. Prove that $\{f_n\}$ converges uniformly on K.

18. Let $\{f_n\}$ be a uniformly bounded sequence of functions which are Riemann-integrable on $[a,b]$ and put

$$F_n(x) = \int_a^x f_n(t)\,\mathrm{d}t$$

for $x \in [a,b]$. Prove that there exists a subsequence $\{F_{n_k}\}$ which converges uniformly on $[a,b]$.

19. Let K be a compact metric space, let S be a subset of $\mathscr{C}(K)$. Prove that S is compact (with respect to the metric defined in Section 7.14) if and only if S is uniformly closed, pointwise bounded, and equicontinuous. (If S is not equicontinuous, then S contains a sequence which has no equicontinuous subsequence, hence has no subsequence that converges uniformly on K.)

20. If f is continuous on $[0,1]$ and if

$$\int_0^1 f(x)x^n\,\mathrm{d}x = 0$$

for $n = 0, 1, 2, \ldots$, prove that $f(x) = 0$ on $[0,1]$. (Hint: The integral of the product of f with any polynomial is zero. Use the Weierstrass theorem to show that $\int_0^1 f^2(x)\,\mathrm{d}x = 0$.)

22. Assume $f \in \mathscr{R}(\alpha)$ on $[a,b]$, and prove that there are polynomials P_n such that

$$\lim_{n\to\infty} \int_a^b |f - P_n|^2\,\mathrm{d}\alpha = 0$$

(Compare with Exercise 6.12.)

23. Put $P_0 = 0$, and define for $n = 0, 1, 2, \ldots$

$$P_{n+1}(x) = P_n(x) + \frac{x^2 - P_n^2(x)}{2}$$

Prove that

$$\lim_{n\to\infty} P_n(x) = |x|$$

uniformly on $[-1,1]$. This makes it possible to prove the Stone-Weierstrass theorem without first proving Theorem 7.26. (Hint: Use the identity

$$|x| - P_{n+1}(x) = [|x| - P_n(x)]\left[1 - \frac{|x| + P_n(x)}{2}\right]$$

to prove that $0 \leq P_n(x) \leq P_{n+1}(x) \leq |x|$ if $|x| \leq 1$ and that

$$|x| - P_n(x) \leq |x| \left(1 - \frac{|x|}{2}\right)^n < \frac{2}{n+1}$$

if $|x| < 1$.)

25. Suppose ϕ is a continuous bounded real function in the strip defined by $0 \leq x \leq 1$, $-\infty < y < \infty$. Prove that the initial-value problem

$$y' = \phi(x, y) \qquad\qquad y(0) = c$$

has a solution. Note that the hypotheses of this existence theorem are less stringent than those of the corresponding uniqueness theorem; see Exercise 5.27. (Hint: Fix n. For $i = 0, \dots, n$, put $x_i = i/n$. Let f_n be a continuous function on $[0, 1]$ such that $f_n(0) = c$, let

$$f_n'(t) = \phi(x_i, f_n(x_i))$$

if $x_i < t < x_{i+1}$, and put

$$\Delta_n(t) = f_n'(t) - \phi(t, f_n(t))$$

except at points x_i, where $\Delta_n(t) = 0$. Then

$$f_n(x) = c + \int_0^x [\phi(t, f_n(t)) + \Delta_n(t)] \, \mathrm{d}t$$

Choose $M < \infty$ so that $|\phi| \leq M$. Verify the following assertions.

(a) $|f_n'| \leq M$, $|\Delta_n| \leq 2M$, $\Delta_n \in \mathscr{R}$, and $|f_n| \leq |c| + M = M_1$ say, on $[0, 1]$, for all n.

(b) $\{f_n\}$ is equicontinuous on $[0, 1]$ since $|f_n'| \leq M$.

(c) Some $\{f_{n_k}\}$ converges to some f, uniformly on $[0, 1]$.

(d) Since ϕ is uniformly continuous on the rectangle $0 \leq x \leq 1$, $|y| \leq M_1$,

$$\phi(t, f_{n_k}(t)) \to \phi(t, f(t))$$

uniformly on $[0, 1]$.

(e) $\Delta_n(t) \to 0$ uniformly on $[0, 1]$ since

$$\Delta_n(t) = \phi(x_i, f_n(x_i)) - \phi(t, f_n(t))$$

in (x_i, x_{i+1}).

(f) Hence

$$f(x) = c + \int_0^x \phi(t, f(t)) \, \mathrm{d}t$$

This f is a solution of the given problem.)

26. Prove an analogous existence theorem for the initial-value problem

$$\mathbf{y}' = \mathbf{\Phi}(\mathbf{x}, \mathbf{y}) \qquad\qquad \mathbf{y}(0) = \mathbf{c}$$

where now $\mathbf{c} \in \mathbb{R}^k$, $\mathbf{y} \in \mathbb{R}^k$, and $\mathbf{\Phi}$ is a continuous bounded mapping of the part of $\mathbb{R}^{k+1}$ defined by $0 \leq x \leq 1$, $\mathbf{y} \in \mathbb{R}^k$ into $\mathbb{R}^k$. Compare Exercise 5.28. (Hint: Use the vector-valued version of Theorem 7.25.)

References

Cunningham, F., Jr. (1967). Taking limits under the integral sign. *Mathematics Magazine*, *40*(4), 179–186. https://doi.org/10.2307/2688673

Fine, N. J. (1966). *The American Mathematical Monthly*, *73*(7), 782. https://doi.org/10.2307/2314000

Kestelman, H. (1970). Riemann integration of limit functions. *The American Mathematical Monthly*, *77*(2), 182–187. https://doi.org/10.2307/2317340

Rudin, W. (1976). *Principles of mathematical analysis* (A. A. Arthur & S. L. Langman, Eds.; Third). McGraw-Hill.

Schoenberg, I. J. (1938). On the Peano curve of Lebesgue. *Bulletin of the American Mathematical Society*, *44*, 519. https://doi.org/10.1090/S0002-9904-1938-06792-4

Treil, S. (2017). *Linear algebra done wrong* [http://www.math.brown.edu/streil/papers/LADW/LADW_2017-09-04.pdf].