COMP3670/6670: Introduction to Machine Learning

Release Date. Aug 4th, 2021
Due Date. 11:59pm, Aug 29th, 2021
Maximum credit. 100

Exercise 1 **Solving Linear Systems** (4+4 credits)

Find the set $\mathcal{S}$ of all solutions $\mathbf{x}$ of the following inhomogenous linear systems $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{A}$ and $\mathbf{b}$ are defined as follows. Write the solution space $\mathcal{S}$ in parametric form.

(a)

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 5 \\ 1 & 4 & 3 \\ 2 & 7 & 1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} -4 \\ -2 \\ -2 \end{bmatrix}$$

Solution. We form the augmented matrix, and row reduce.

$$\left[\begin{array}{ccc|c} 0 & 1 & 5 & -4 \\ 1 & 4 & 3 & -2 \\ 2 & 7 & 1 & -2 \end{array}\right]$$

$$\Big\downarrow \text{Swap } R_3 \text{ and } R_1.$$

$$\left[\begin{array}{ccc|c} 2 & 7 & 1 & -2 \\ 1 & 4 & 3 & -2 \\ 0 & 1 & 5 & -4 \end{array}\right]$$

$$\Big\downarrow R_1 := R_1 - 2R_2$$

$$\left[\begin{array}{ccc|c} 0 & -1 & -5 & 2 \\ 1 & 4 & 3 & -2 \\ 0 & 1 & 5 & -4 \end{array}\right]$$

$$\Big\downarrow R_1 := R_1 + R_3$$

$$\left[\begin{array}{ccc|c} 0 & 0 & 0 & -2 \\ 1 & 4 & 3 & -2 \\ 0 & 1 & 5 & -4 \end{array}\right]$$

The first line of the matrix gives $0 = -2$, a contradiction. No solutions exist, $\mathcal{S} = \varnothing$.

(b)

$$\mathbf{A} = \begin{bmatrix} 2 & 3 & 1 \\ 4 & 0 & 3 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 6 \\ 12 \end{bmatrix}$$

Solution.

$$\mathcal{S} = \left\{ \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} -3/4 \\ 1/6 \\ 1 \end{bmatrix} x_3 : x_3 \in \mathbb{R} \right\}$$

Solution. We form the augmented matrix, and row reduce.

$$\left[\begin{array}{ccc|c} 2 & 3 & 1 & 6 \\ 4 & 0 & 3 & 12 \end{array}\right]$$

$$\Big\downarrow R_2 := R_2 - 2R_1$$

$$\left[\begin{array}{ccc|c} 2 & 3 & 1 & 6 \\ 0 & -6 & 1 & 0 \end{array}\right]$$

$$\Big\downarrow R_1 := R_1 + \frac{1}{2}R_2$$

$$\left[\begin{array}{ccc|c} 2 & 0 & 3/2 & 6 \\ 0 & -6 & 1 & 0 \end{array}\right]$$

$$\Big\downarrow \begin{array}{l} R_1 := \frac{1}{2}R_1 \\ R_2 := -\frac{1}{6}R_2 \end{array}$$

$$\left[\begin{array}{ccc|c} 1 & 0 & 3/4 & 3 \\ 0 & 1 & -1/6 & 0 \end{array}\right]$$

At this point we can read off the equations $x_1 + 3/4x_3 = 3$ and $x_2 - \frac{1}{6}x_3 = 0$. Rearranging the second gives $x_2 = x_3/6$, and rearranging the first gives $x_1 = 3 - \frac{3}{4}x_3$. Here, x_3 is a free variable. So, the solution space is given as

$$\mathcal{S} = \left\{ \begin{bmatrix} 3 - \frac{3}{4}x_3 \\ x_3/6 \\ x_3 \end{bmatrix} : x_3 \in \mathbb{R} \right\} = \left\{ \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} -3/4 \\ 1/6 \\ 1 \end{bmatrix} x_3 : x_3 \in \mathbb{R} \right\}$$

Exercise 2 **Inverses** (4 credits)

For what values of $[a, b, c]^T \in \mathbb{R}^3$ does the inverse of the following matrix exist?

$$\begin{bmatrix} 1 & a & b \\ 1 & 1 & c \\ 1 & 1 & 1 \end{bmatrix}$$

Solution. We compute the determinant using Sarrus' Rule, and obtain

$$1 + ac + b - b - c - a = ac - a - c + 1 = (a-1)(c-1)$$

Since the matrix is invertable iff the determinant is non-zero, we have that $(a-1)(c-1) \neq 0$, so $a \neq 1$ and $c \neq 1$. So, the matrix is invertable for all $[a, b, c]^T \in \mathbb{R}^3$ such that both $a \neq 1$ and $c \neq 1$.

Exercise 3 **Subspaces** (3+3+3+3 credits)

Which of the following sets are subspaces of $\mathbb{R}^3$? Prove your answer. (That is, if it is a subspace, you must demonstrate the subspace axioms are satisfied, and if it is not a subspace, you must show which axiom fails.)

(a) $A = \{(x, y) \in \mathbb{R}^2 : x \geq 0, y \geq 0\}$

Solution. NO, fails closure under scalar multiplication.
$[1, 1]^T \in A$ but $-1 \cdot [1, 1]^T = [-1, -1] \notin A$.

(b) $B = \{(x, y, z) : x + y + z = 0\}$.

Solution. YES, we check the requisite three properties.

(a) Closure under scalar multiplication.
Let $\mathbf{x} \in B$. Then $\mathbf{x} = [x, y, z]^T$ with $x + y + z = 0$.
Then $c\mathbf{x} = [cx, cy, cz]^T$, and $cx + cy + cz = c(x + y + z) = c \cdot 0 = 0$, so $c\mathbf{x} \in B$.

(b) Closure under vector addition.
Let $\mathbf{x}, \mathbf{y} \in B$. Then $\mathbf{x} = [x_1, x_2, x_3]^T$ with $x_1 + x_2 + x_3 = 0$ and $\mathbf{y} = [y_1, y_2, y_3]^T$ with $y_1 + y_2 + y_3 = 0$. Then $\mathbf{x} + \mathbf{y} = [x_1 + y_1, x_2 + y_2, x_3 + y_3]^T$, and $(x_1 + y_1) + (x_2 + y_2) + (x_3 + y_3) = (x_1 + x_2 + x_3) + (y_1 + y_2 + y_3) = 0 + 0 = 0$, so $\mathbf{x} + \mathbf{y} \in B$.

(c) Contains the zero vector.
Clearly $\mathbf{0} \in B$, as $\mathbf{0} = [0, 0, 0]^T$ and $0 + 0 + 0 = 0$.

(c) $C = \{(x, y) \in \mathbb{R}^2 : x = 0 \text{ or } y = 0\}$

Solution. NO, fails closure under addition. We have that $[0, 1]^T \in C$ and $[1, 0]^T \in C$, but $[0, 1]^T + [1, 0]^T = [1, 1]^T \notin C$.

(d) $D =$ The set of all solutions $\mathbf{x}$ to the matrix equation $\mathbf{Ax} = \mathbf{b}$, for some matrix $\mathbf{A}$ and some vector $\mathbf{b}$. (Hint: Your answer may depend on $\mathbf{A}$ and $\mathbf{b}$.)

Solution. Yes if and only if $\mathbf{b} = \mathbf{0}$.
We check the three axioms.

(a) Closure under scalar multiplication.
Let $\mathbf{x} \in D$. Then $\mathbf{Ax} = \mathbf{b}$. We have that $\mathbf{A}(\lambda\mathbf{x}) = \lambda(\mathbf{Ax}) = \lambda\mathbf{b}$. For $\lambda\mathbf{b} = \mathbf{b}$ for any choice of λ, it must be the case that $\mathbf{b} = \mathbf{0}$. So $\lambda\mathbf{x} \in D$ conditional on $\mathbf{b} = \mathbf{0}$.

(b) Closure under vector addition.
Let $\mathbf{x}, \mathbf{y} \in D$. Then $\mathbf{Ax} = \mathbf{b}$ and $\mathbf{Ay} = \mathbf{b}$. But then $\mathbf{A}(\mathbf{x} + \mathbf{y}) = \mathbf{Ax} + \mathbf{Ay} = \mathbf{b} + \mathbf{b} = 2\mathbf{b}$. Now, $\mathbf{b} = 2\mathbf{b}$ is true if and only if $\mathbf{b} = \mathbf{0}$, so we have closure under addition conditional on $\mathbf{b} = \mathbf{0}$.

(c) Contains the zero vector.
$\mathbf{A0} = \mathbf{b}$ is true if and only if $\mathbf{b} = \mathbf{0}$, so this axiom is also conditional on $\mathbf{b} = \mathbf{0}$.

To conclude, the three axioms hold if $\mathbf{b} = \mathbf{0}$, and all of them don't if $\mathbf{b} \neq \mathbf{0}$.

Exercise 4 **Linear Independence** (4+8+8 credits)

Let V and W be vector spaces. Let $T : V \to W$ be a linear transformation.

(a) Prove that $T(\mathbf{0}) = \mathbf{0}$.

Solution. $T(\mathbf{0}) = T(\mathbf{0} + \mathbf{0}) = T(\mathbf{0}) + T(\mathbf{0})$. Since $T(\mathbf{0}) = T(\mathbf{0}) + T(\mathbf{0})$, we subtract $T(\mathbf{0})$ from both sides to obtain $\mathbf{0} = T(\mathbf{0})$, as required.

(b) For any integer $n \geq 1$, prove that given a set of vectors $\{\mathbf{v}_1, \ldots \mathbf{v}_n\}$ in V and a set of coefficients $\{c_1, \ldots, c_n\}$ in $\mathbb{R}$, that

$$T(c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n) = c_1 T(\mathbf{v}_1) + \ldots + c_n T(\mathbf{v}_n)$$

Solution. We proceed by induction. The base case follows immediately from the definition of linearity of T, as $T(c_1\mathbf{v}_1) = c_1 T(\mathbf{v}_1)$. Step case, assume that

$$T(c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n) = c_1 T(\mathbf{v}_1) + \ldots + c_n T(\mathbf{v}_n) \qquad \text{(Induction Hypothesis)}$$

for any $\{\mathbf{v}_1, \ldots \mathbf{v}_n\}$ in V, $\{c_1, \ldots, c_n\}$ in $\mathbb{R}$. We now prove that

$$T(c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n + c_{n+1}\mathbf{v}_{n+1}) = c_1 T(\mathbf{v}_1) + \ldots + c_n T(\mathbf{v}_n) + c_{n+1} T(\mathbf{v}_{n+1})$$

for any $\{\mathbf{v}_1, \ldots \mathbf{v}_{n+1}\}$ in V, $\{c_1, \ldots, c_{n+1}\}$ in $\mathbb{R}$.

$$\begin{aligned}
&T(c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n + c_{n+1}\mathbf{v}_{n+1}) && \\
=&T((c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n) + c_{n+1}\mathbf{v}_{n+1}) && \text{Vector addition is associative} \\
=&T(c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n) + T(c_{n+1}\mathbf{v}_{n+1}) && T \text{ distributes over vector addition} \\
=&c_1T(\mathbf{v}_1) + \ldots + c_nT(\mathbf{v}_n) + T(c_{n+1}\mathbf{v}_{n+1}) && \text{Induction Hypothesis} \\
=&c_1T(\mathbf{v}_1) + \ldots + c_nT(\mathbf{v}_n) + c_{n+1}T(\mathbf{v}_{n+1}) && T \text{ distributes over scalar multiplication}
\end{aligned}$$

as required.

(c) Let $\{\mathbf{v}_1, \ldots \mathbf{v}_n\}$ be a set of linearly **dependent** vectors in V.
Define $\mathbf{w}_1 := T(\mathbf{v}_1), \ldots, \mathbf{w}_n := T(\mathbf{v}_n)$.
Prove that $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ is a set of linearly **dependent** vectors in W.

Solution. We are given that $\{\mathbf{v}_1, \ldots \mathbf{v}_n\}$ is a set of linearly dependant vectors in V. Then there exists non-trivial[1] solutions to the equation

$$c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n = \mathbf{0}$$

Apply the transformation T to both sides of the equation,

$$\begin{aligned}
T(c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n) &= T(\mathbf{0}) && \\
T(c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n) &= \mathbf{0} && \text{Exercise 4a} \\
c_1T(\mathbf{v}_1) + \ldots + c_nT(\mathbf{v}_n) &= \mathbf{0} && \text{Exercise 4b} \\
c_1\mathbf{w}_1 + \ldots + c_n\mathbf{w}_n &= \mathbf{0} && \text{Definition of } \mathbf{w}_1, \ldots, \mathbf{w}_n
\end{aligned}$$

and hence $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ is a set of linearly dependant vectors in W.

Exercise 5 **Inner Products** (4+8 credits)

(a) Show that if an inner product $\langle \cdot, \cdot \rangle$ is symmetric and linear in the first argument, then it is bilinear.

Solution. Suppose that $\langle \cdot, \cdot \rangle$ is a symmetric and linear in the first argument inner product. Then,

$$\langle \mathbf{x}, a\mathbf{y} + b\mathbf{z} \rangle = \langle a\mathbf{y} + b\mathbf{z}, \mathbf{x} \rangle = a \langle \mathbf{y}, \mathbf{x} \rangle + b \langle \mathbf{z}, \mathbf{x} \rangle = a \langle \mathbf{x}, \mathbf{y} \rangle + b \langle \mathbf{x}, \mathbf{z} \rangle$$

Hence $\langle \cdot, \cdot \rangle$ is linear in the second argument, and hence bilinear.

(b) Define $\langle \cdot, \cdot \rangle$ for all $\mathbf{x} = [x_1, x_2]^T \in \mathbb{R}^2$ and $\mathbf{y} = [y_1, y_2]^T \in \mathbb{R}^2$ as

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1y_1 + x_2y_2 + 2(x_1y_2 + x_2y_1)$$

Which of the three inner product axioms does $\langle \cdot, \cdot \rangle$ satisfy?

Solution. Symmetry and bilinearity are satisfied, but positive definiteness is not.

(a) Symmetry. We verify that $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$.

$$\begin{aligned}
\langle \mathbf{x}, \mathbf{y} \rangle =& x_1y_1 + x_2y_2 + 2(x_1y_2 + x_2y_1) \\
=& y_1x_1 + y_2x_2 + 2(x_2y_1 + x_1y_2) \\
=& \langle \mathbf{y}, \mathbf{x} \rangle
\end{aligned}$$

(b) Bilinearity. We first verify linearity in the first argument, that is, that $\langle a\mathbf{x} + b\mathbf{y}, \mathbf{z} \rangle = a \langle \mathbf{x}, \mathbf{z} \rangle + b \langle \mathbf{y}, \mathbf{z} \rangle$ for all $a, b \in \mathbb{R}$ and all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^2$.

$$\begin{aligned}
\langle a\mathbf{x} + b\mathbf{y}, \mathbf{z} \rangle &= (a\mathbf{x} + b\mathbf{y})_1z_1 + (a\mathbf{x} + b\mathbf{y})_2z_2 + 2((a\mathbf{x} + b\mathbf{y})_1z_2 + (a\mathbf{x} + b\mathbf{y})_2z_1 \\
&= ax_1z_1 + by_1z_1 + ax_2z_2 + by_2z_2 + 2ax_1z_2 + 2by_1z_2 + 2ax_2z_1 + 2by_2z_1 \\
&= (ax_1z_1 + ax_2z_2 + 2ax_1z_2 + 2ax_2z_1) + (by_1z_1 + by_2z_2 + 2by_1z_2 + 2by_2z_1) \\
&= a(x_1z_1 + x_2z_2 + 2(x_1z_2 + x_2z_1)) + b(y_1z_1 + y_2z_2 + 2(y_1z_2 + 2y_2z_1)) \\
&= a\langle \mathbf{x}, \mathbf{z} \rangle + b\langle \mathbf{y}, \mathbf{z} \rangle
\end{aligned}$$

By Exercise 5a, symmetry and linearity in the first argument imply bilinearity.

[1] That is, solutions other than $c_1 = c_2 = \ldots = c_n = 0$.

(c) Positive Definiteness. This propery fails, choose $\mathbf{x} = [1, -1]^T$. Then

$$\langle \mathbf{x}, \mathbf{x} \rangle = x_1^2 + x_2^2 + 4x_1x_2 = 1^2 + (-1)^2 + 4(1)(-1) = 1 + 1 - 4 = -2 \ngeq 0$$

Exercise 6 **Orthogonality** (8+6 credits)

Let V denote a vector space together with an inner product $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$.

Let $\mathbf{x}, \mathbf{y}$ be **non-zero** vectors in V.

(a) Prove or disprove that if $\mathbf{x}$ and $\mathbf{y}$ are orthogonal, then they are linearly independent.

Solution. The statement is true. We are given that $\mathbf{x}$ and $\mathbf{y}$ are orthogonal, so $\langle \mathbf{x}, \mathbf{y} \rangle = 0$. Assume for a contradiction that $\mathbf{x}$ and $\mathbf{y}$ are linearly dependant, so there exists non-trivial solutions to the equation

$$c_1\mathbf{x} + c_2\mathbf{y} = \mathbf{0}.$$

Since the solution is non trivial, at least one of the c_i is non-zero. Proceed by cases.
Case 1: $c_1 \neq 0$.
Then we inner product both sides with $\mathbf{x}$,

$$\begin{aligned} \langle \mathbf{c_1x} + \mathbf{c_2y}, \mathbf{x} \rangle &= \langle \mathbf{0}, \mathbf{x} \rangle && \\ \langle \mathbf{c_1x} + \mathbf{c_2y}, \mathbf{x} \rangle &= 0 && \text{Tutorial 2} \\ c_1\langle \mathbf{x}, \mathbf{x} \rangle + c_2\langle \mathbf{y}, \mathbf{x} \rangle &= 0 && \text{Bilinearity} \\ c_1\langle \mathbf{x}, \mathbf{x} \rangle &= 0 && \text{Orthogonality of } \mathbf{x} \text{ and } \mathbf{y} \end{aligned}$$

Now, since $c_1 \neq 0$, we have that $\langle \mathbf{x}, \mathbf{x} \rangle = 0$, and then by positive definiteness, $\mathbf{x} = \mathbf{0}$, a contradiction.
Case 2: $c_2 \neq 0$.
Then we inner product both sides with $\mathbf{y}$,

$$\begin{aligned} \langle \mathbf{c_1x} + \mathbf{c_2y}, \mathbf{y} \rangle &= \langle \mathbf{0}, \mathbf{y} \rangle && \\ \langle \mathbf{c_1x} + \mathbf{c_2y}, \mathbf{y} \rangle &= 0 && \text{Tutorial 2} \\ c_1\langle \mathbf{x}, \mathbf{y} \rangle + c_2\langle \mathbf{y}, \mathbf{y} \rangle &= 0 && \text{Bilinearity} \\ c_2\langle \mathbf{y}, \mathbf{y} \rangle &= 0 && \text{Orthogonality of } \mathbf{x} \text{ and } \mathbf{y} \end{aligned}$$

Now, since $c_2 \neq 0$, we have that $\langle \mathbf{y}, \mathbf{y} \rangle = 0$, and then by positive definiteness, $\mathbf{y} = \mathbf{0}$, a contradiction.
So, in either case we get a contradiction, and hence there are no non-trivial solutions to $c_1\mathbf{x}+c_2\mathbf{y} = \mathbf{0}$.
We conclude that $\mathbf{x}$ and $\mathbf{y}$ are linearly independent.

(b) Prove or disprove that if $\mathbf{x}$ and $\mathbf{y}$ are linearly independent, then they are orthogonal.

Solution. No. For a counter example, choose the vector space $V = \mathbb{R}^2$ equipped with the standard Eucledian dot product. Let $\mathbf{x} = (0, 1)^T$, $\mathbf{y} = (1, 1)^T$. They are linearly independent, as by solving $c_1\mathbf{x} + c_2\mathbf{y} = \mathbf{0}$ for c_1, c_2, we recover the two equations $0c_1 + 1c_2 = 0$ and $1c_1 + 1c_2 = 0$. The first equation gives $c_1 = 0$, substituting into the second gives $c_2 = 0$, so $\mathbf{x}, \mathbf{y}$ are linearly independant. But

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x} \cdot \mathbf{y} = 0 \cdot 1 + 1 \cdot 1 = 1 \neq 0$$

so they are not orthogonal.

Exercise 7 **Properties of Norms** (4+4+10 credits)

Given a vector space V with two norms $\|\cdot\|_a : V \to \mathbb{R}_{\geq 0}$ and $\|\cdot\|_b : V \to \mathbb{R}_{\geq 0}$, we say that the two norms $\|\cdot\|_a$ and $\|\cdot\|_b$ are ε-*equivalent* if for any $\mathbf{v} \in V$, we have that

$$\varepsilon\|\mathbf{v}\|_a \leq \|\mathbf{v}\|_b \leq \frac{1}{\varepsilon}\|\mathbf{v}\|_a.$$

where $\varepsilon \in (0, 1]$.

If $\|\cdot\|_a$ is ε-equivalent to $\|\cdot\|_b$, we denote this as $\|\cdot\|_a \overset{\varepsilon}{\sim} \|\cdot\|_b$.

(a) Is ε-equivalence reflexive for all $\varepsilon \in (0,1]$?
(Is it true that $\|\cdot\|_a \overset{\varepsilon}{\sim} \|\cdot\|_a$?)

Solution. Yes. Clearly $\varepsilon \leq 1 \leq \frac{1}{\varepsilon}$. Then

$$\varepsilon\|\cdot\|_a \leq \|\cdot\|_a \leq \frac{1}{\varepsilon}\|\cdot\|_a$$

(noting that norms are always non-negative.)

(b) Is ε-equivalence symmetric for all $\varepsilon \in (0,1]$?
(Does $\|\cdot\|_a \overset{\varepsilon}{\sim} \|\cdot\|_b$ imply $\|\cdot\|_b \overset{\varepsilon}{\sim} \|\cdot\|_a$?)

Solution. Yes, as if

$$\varepsilon\|\cdot\|_a \leq \|\cdot\|_b \leq \frac{1}{\varepsilon}\|\cdot\|_a$$

then $\varepsilon\|\cdot\|_a \leq \|\cdot\|_b$ implies $\|\cdot\|_a \leq \frac{1}{\varepsilon}\|\cdot\|_b$ and $\|\cdot\|_b \leq \frac{1}{\varepsilon}\|\cdot\|_a$ implies $\varepsilon\|\cdot\|_b \leq \|\cdot\|_a$. Therefore,

$$\varepsilon\|\cdot\|_b \leq \|\cdot\|_a \leq \frac{1}{\varepsilon}\|\cdot\|_b$$

(c) Assuming that $V = \mathbb{R}^2$, prove that $\|\cdot\|_1 \overset{\varepsilon}{\sim} \|\cdot\|_2$. for the largest ε possible.

Solution.

$$\|\mathbf{v}\|_1^2 = (|v_1| + |v_2|)^2 = |v_1|^2 + 2|v_1||v_2| + |v_2|^2 \geq |v_1|^2 + |v_2|^2 = \|\mathbf{v}\|_2^2$$

$$\therefore \|\mathbf{v}\|_2 \leq \|\mathbf{v}\|_1$$

For the other case, first consider the fact that $(x-y)^2 \geq 0$, which means $x^2 - 2xy + y^2 \geq 0$ and $2xy \leq x^2 + y^2$. Then,

$$\|\mathbf{v}\|_1^2 = |v_1|^2 + 2|v_1||v_2| + |v_2|^2 \leq |v_1|^2 + (|v_1|^2 + |v_2|^2) + |v_2|^2 = 2\|\mathbf{v}\|_2^2$$

$$\therefore \|\mathbf{v}\|_1 \leq \sqrt{2}\|\mathbf{v}\|_2$$

Putting the two results together,

$$\|\mathbf{v}\|_2 \leq \|\mathbf{v}\|_1 \leq \sqrt{2}\|\mathbf{v}\|_2$$

This means that any choice of ε such that $1/\varepsilon \geq \sqrt{2}$ is valid, or any $\varepsilon \leq 1/\sqrt{2}$, so we choose $\varepsilon = 1/\sqrt{2}$. Now, to demonstrate that any choice of $\varepsilon > 1/\sqrt{2}$ is invalid, consider $\mathbf{x} = [1,1]^T$. Then $\|\mathbf{x}\|_1 = 2$ and $\|\mathbf{x}\|_2 = \sqrt{2}$, giving us

$$\|\mathbf{x}\|_1 \leq \frac{1}{\varepsilon}\|\mathbf{x}\|_2$$

$$2 \leq \frac{1}{\varepsilon}\sqrt{2}$$

$$2/\sqrt{2} \leq \frac{1}{\varepsilon}$$

$$\sqrt{2}/2 \geq \varepsilon$$

$$\varepsilon \leq \sqrt{2}/2 = 1/\sqrt{2}$$

Hence we choose $\varepsilon = 1/\sqrt{2}$, the largest possible value such that $\|\cdot\|_1 \overset{\varepsilon}{\sim} \|\cdot\|_2$.

Exercise 8 **Projections** (3+3+3+3 credits)

Consider the Euclidean vector space $\mathbb{R}^3$ with the dot product. A subspace $U \subset \mathbb{R}^3$ and vector $\mathbf{x} \in \mathbb{R}^3$ are given by

$$U = \text{span}\left\{\begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}2\\1\\0\end{bmatrix}\right\}, \mathbf{x} = \begin{bmatrix}12\\12\\18\end{bmatrix}$$

(a) Show that $\mathbf{x} \notin U$.

Solution. We need to show there is no solution in $c_1, c_2 \in \mathbb{R}$ for

$$c_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} = \mathbf{x}$$

We form the augmented matrix, and solve,

$$\left[\begin{array}{cc|c} 1 & 2 & 12 \\ 1 & 1 & 12 \\ 1 & 0 & 18 \end{array}\right] \sim \left[\begin{array}{cc|c} 0 & 2 & -6 \\ 0 & 1 & -6 \\ 1 & 0 & 18 \end{array}\right] \sim \left[\begin{array}{cc|c} 0 & 0 & 6 \\ 0 & 1 & -6 \\ 1 & 0 & 18 \end{array}\right]$$

at which point it is clear that no solution can exist, as the first equation gives $0 = 6$. Hence, we cannot write $\mathbf{x}$ as a linear combination of the vectors that span U, and thus $\mathbf{x} \notin U$.

(b) Determine the orthogonal projection of $\mathbf{x}$ onto U, denoted $\pi_U(\mathbf{x})$.

Solution. We let

$$\mathbf{B} = \begin{bmatrix} 1 & 2 \\ 1 & 1 \\ 1 & 0 \end{bmatrix}$$

and compute the projection matrix $P_\pi = \mathbf{B}(\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T$

$$P_\pi = \frac{1}{6} \begin{bmatrix} 5 & 2 & -1 \\ 2 & 2 & 2 \\ -1 & 2 & 5 \end{bmatrix}$$

Now, for the projection, we compute $P_\pi \mathbf{x}$.

$$\pi_U(\mathbf{x}) = P_\pi \mathbf{x} = \frac{1}{6} \begin{bmatrix} 5 & 2 & -1 \\ 2 & 2 & 2 \\ -1 & 2 & 5 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 11 \\ 14 \\ 17 \end{bmatrix}$$

(c) Show that $\pi_U(\mathbf{x})$ can be written as a linear combination of $[1, 1, 1]^T$ and $[2, 2, 3]^T$.

Solution. We verify that there exists c_1, c_2 such that

$$c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 = P_\pi \mathbf{x}$$

by forming the augmented matrix, and solving for c_1, c_2,

$$\left[\begin{array}{cc|c} 1 & 2 & 11 \\ 1 & 1 & 14 \\ 1 & 0 & 17 \end{array}\right] \sim \left[\begin{array}{cc|c} 0 & 2 & -6 \\ 0 & 1 & -3 \\ 1 & 0 & 17 \end{array}\right] \sim \left[\begin{array}{cc|c} 0 & 0 & 0 \\ 0 & 1 & -3 \\ 1 & 0 & 17 \end{array}\right]$$

giving $c_1 = 17, c_2 = -3$.

(d) Determine the distance $d(\mathbf{x}, U) := \min_{\mathbf{y} \in U} \|\mathbf{x} - \mathbf{y}\|_2$.

Solution. The vector in U closest to $\mathbf{x}$ is the orthogonal projection $\pi_U(x)$, computed above. Hence,

$$d(\mathbf{x}, U) = \|\mathbf{x} - \pi_U(\mathbf{x})\|_2 = \left\| \begin{bmatrix} 12 \\ 12 \\ 18 \end{bmatrix} - \begin{bmatrix} 11 \\ 14 \\ 17 \end{bmatrix} \right\|_2 = \left\| \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \right\|_2 = \sqrt{1^2 + 2^2 + 1^2} = \sqrt{6}$$