COMP3670 2021 Theory Assignment 1

Yuxuan Lin – u6828533@anu.edu.au

August 24, 2021

Introduction to Machine Learning

By turning in this assignment, I agree by the ANU honor code and declare that all of this is my own work.

Exercise 1

(a) We first perform Gaussian Elimination on the augmented matrix $\left[\begin{array}{ccc|c} 0 & 1 & 5 & -4 \\ 1 & 4 & 3 & -2 \\ 2 & 7 & 1 & -2 \end{array}\right]$ to get the row-echelon form.

$$\begin{bmatrix} 0 & 1 & 5 & -4 \\ 1 & 4 & 3 & -2 \\ 2 & 7 & 1 & -2 \end{bmatrix} \xrightarrow[R_1,R_2]{Swap} \begin{bmatrix} 1 & 4 & 3 & -2 \\ 0 & 1 & 5 & -4 \\ 2 & 7 & 1 & -2 \end{bmatrix} \xrightarrow[R_2,R_3]{Swap} \begin{bmatrix} 1 & 4 & 3 & -2 \\ 2 & 7 & 1 & -2 \\ 0 & 1 & 5 & -4 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 4 & 3 & -2 \\ 2 & 7 & 1 & -2 \\ 0 & 1 & 5 & -4 \end{bmatrix} \xrightarrow{R_2=R_2-2R_1} \begin{bmatrix} 1 & 4 & 3 & -2 \\ 0 & -1 & -5 & 2 \\ 0 & 1 & 5 & -4 \end{bmatrix} \xrightarrow{R_3=R_2-R_3} \begin{bmatrix} 1 & 4 & 3 & -2 \\ 0 & -1 & -5 & 2 \\ 0 & 0 & 0 & -2 \end{bmatrix}$$

We can get the REF $\left[\begin{array}{ccc|c} 1 & 4 & 3 & -2 \\ 0 & -1 & -5 & 2 \\ 0 & 0 & 0 & -2 \end{array}\right]$. From the third row we have $0+0+0=-2$, which is a contradiction, hence no solution exists.

The solution space of $\boldsymbol{Ax} = \boldsymbol{b}$ is : $\mathcal{S} = \emptyset$.

(b) We first perform Gaussian Elimination on the augmented matrix $\left[\begin{array}{ccc|c} 2 & 3 & 1 & 6 \\ 4 & 0 & 3 & 12 \end{array}\right]$ to get the row-echelon form.

$$\begin{bmatrix} 2 & 3 & 1 & 6 \\ 4 & 0 & 3 & 12 \end{bmatrix} \xrightarrow{R_2=R_2-2R_1} \begin{bmatrix} 2 & 3 & 1 & 6 \\ 0 & -6 & 1 & 0 \end{bmatrix}$$

We can get the REF $\begin{bmatrix} 2 & 3 & 1 & \bigm| & 6 \\ 0 & -6 & 1 & \bigm| & 0 \end{bmatrix}$. We observe that the third column is not a pivot column, thus x_3 is a free variable. Then, we perform the following steps to get a general solution to $\boldsymbol{Ax} = \boldsymbol{b}$.

Step 1: Find a particular solution to $\boldsymbol{Ax} = \boldsymbol{b}$.

Let the only free variable be 0 (i.e. let $x_3 = 0$), calculate the value of basic variables.

$$\begin{bmatrix} 2 & 3 & 1 & \bigm| & 6 \\ 0 & -6 & 1 & \bigm| & 0 \end{bmatrix}$$
$$0 - 6x_2 + 0 = 0 \implies x_2 = 0$$
$$2x_1 + 3x_2 + 0 = 6 \implies x_1 = 3$$

Our particular solution to $\boldsymbol{Ax} = \boldsymbol{b}$ is : $\begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}$.

Step 2: Find all solutions to $\boldsymbol{Ax} = \boldsymbol{0}$.

Let the only free variable to be 1 (i.e. let $x_3 = 1$), calculate the value of basic variables.

$$\begin{bmatrix} 2 & 3 & 1 & \bigm| & 0 \\ 0 & -6 & 1 & \bigm| & 0 \end{bmatrix}$$
$$0 - 6x_2 + 1 = 0 \implies x_2 = \frac{1}{6}$$
$$2x_1 + 3x_2 + 1 = 0 \implies x_1 = -\frac{3}{4}$$
$$\begin{bmatrix} -\frac{3}{4} \\ \frac{1}{6} \\ 1 \end{bmatrix} = \gamma \begin{bmatrix} -9 \\ 2 \\ 12 \end{bmatrix}, \gamma \in \mathbb{R}$$

All solutions to $\boldsymbol{Ax} = \boldsymbol{0}$ is : $\left\{ x \in \mathbb{R}^3 : x = \lambda \begin{bmatrix} -9 \\ 2 \\ 12 \end{bmatrix}, \lambda \in \mathbb{R} \right\}$.

Step 3: Combine the solutions from **Step 1** and **Step 2** to the general solution.

The solution space of $\boldsymbol{Ax} = \boldsymbol{b}$ is : $\mathcal{S} = \left\{ x \in \mathbb{R}^3 : x = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix} + \lambda \begin{bmatrix} -9 \\ 2 \\ 12 \end{bmatrix}, \lambda \in \mathbb{R} \right\}$.

Exercise 2

We first perform Gaussian Elimination on the given matrix to get the row-echelon form.

$$\boldsymbol{A} = \begin{bmatrix} 1 & a & b \\ 1 & 1 & c \\ 1 & 1 & 1 \end{bmatrix} \xrightarrow[R_1,R_3]{Swap} \begin{bmatrix} 1 & 1 & 1 \\ 1 & a & b \\ 1 & 1 & c \end{bmatrix} \xrightarrow[R_3=R_3-R_1]{R_2=R_2-R_1} \begin{bmatrix} 1 & 1 & 1 \\ 0 & a-1 & b-1 \\ 0 & 0 & c-1 \end{bmatrix}$$

For an upper triangular matrix $\boldsymbol{A} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & a-1 & b-1 \\ 0 & 0 & c-1 \end{bmatrix}$, its determinant $\det(\boldsymbol{A})$ is equal to the product of all the diagonal elements of A, that is

$$\det(\boldsymbol{A}) = \prod_{i=1}^{3} a_{ii} = 1(a-1)(c-1) = (a-1)(c-1)$$

For the matrix $\boldsymbol{A}$ to be invertible, its determinant should not be zero, that is

$$\det(\boldsymbol{A}) \neq 0 \implies (a-1)(c-1) \neq 0 \implies a \neq 1 \land c \neq 1$$

Hence, the inverse of the given matrix exists if and only if $a \neq 1$ and $c \neq 1$ for the values of $[a, b, c]^T \in \mathbb{R}^3$.

Exercise 3

(a) No, $\boldsymbol{A}$ is not a subspace of $\mathbb{R}^2$. Consider the vector $\boldsymbol{T} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \in \boldsymbol{A}$, its multiplication by the scalar $\lambda = -1$ is : $\lambda \boldsymbol{T} = \begin{bmatrix} -1 \\ -1 \end{bmatrix} \notin \boldsymbol{A}$. Thus, it does not satisfy axiom of closure.

(b) Yes, $\boldsymbol{B}$ is a subspace of $\mathbb{R}^3$.

First we have $\boldsymbol{B} \subseteq \mathbb{R}^3$, $\boldsymbol{B} \neq \emptyset$, and in particular, since $0 + 0 + 0 = 0$, for the zero vector of the vector space $\mathbb{R}^3$ we have $\mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \in \boldsymbol{B}$.

Second, with respect to the outer operation: $\forall \lambda \in \mathbb{R}, \forall \boldsymbol{T} \in \boldsymbol{B}$, $\boldsymbol{T} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, then $x + y + z = 0$, and $\lambda x + \lambda y + \lambda z = \lambda(x + y + z) = 0$, we get $\lambda \boldsymbol{T} = \begin{bmatrix} \lambda x \\ \lambda y \\ \lambda z \end{bmatrix} \in \boldsymbol{B}$.

Third, with respect to the inner operation: $\forall \boldsymbol{T}_1, \boldsymbol{T}_2 \in \boldsymbol{B}$, $\boldsymbol{T}_1 = \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}$, $\boldsymbol{T}_2 = \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix}$, then $x_1+y_1+z_1=0$, $x_2+y_2+z_2=0$, and $(x_1+x_2)+(y_1+y_2)+(z_1+z_2) = (x_1+y_1+z_1)+(x_2+y_2+z_2)=0$, we get $\boldsymbol{T}_1+\boldsymbol{T}_2 = \begin{bmatrix} x_1+x_2 \\ y_1+y_2 \\ z_1+z_2 \end{bmatrix} \in \boldsymbol{B}$.

Thus, it also satisfies axiom of closure. In sum, all the subspace axioms are satisfied, and we have proved that $\boldsymbol{B}$ is a subspace of $\mathbb{R}^3$.

(c) No, $\boldsymbol{C}$ is not a subspace of $\mathbb{R}^2$. Consider the vectors $\boldsymbol{T}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \in \boldsymbol{C}$, $\boldsymbol{T}_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \in \boldsymbol{C}$, the vector of their addition is : $\boldsymbol{T}_1+\boldsymbol{T}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \notin \boldsymbol{C}$. Thus, it does not satisfy axiom of closure.

(d) It depends on the value of $\boldsymbol{b}$.

If $\boldsymbol{b} \neq \mathbf{0}$, $\boldsymbol{D}$ is not a subspace of $\mathbb{R}^n$:

$\forall \boldsymbol{T} \in \boldsymbol{D}$, $\boldsymbol{AT} = \boldsymbol{b} \neq \mathbf{0}$, then $\forall \boldsymbol{T} \in \boldsymbol{D}, \boldsymbol{T} \neq \mathbf{0}$, as known as $\mathbf{0} \notin \boldsymbol{D}$. Thus, it does not satisfy axiom of zero vector.

If $\boldsymbol{b} = \mathbf{0}$, $\boldsymbol{D}$ is a subspace of $\mathbb{R}^n$:

First we have $\boldsymbol{D} \subseteq \mathbb{R}^n$, $\boldsymbol{D} \neq \emptyset$, and in particular, for the zero vector $\mathbf{0}$, we have $\boldsymbol{A} * \mathbf{0} = \mathbf{0} = \boldsymbol{b}$, hence $\mathbf{0} \in \boldsymbol{D}$.

Second, with respect to the outer operation: $\forall \lambda \in \mathbb{R}, \forall \boldsymbol{T} \in \boldsymbol{D}$, then $\boldsymbol{AT} = \boldsymbol{b} = \mathbf{0}$, and $\boldsymbol{A}(\lambda \boldsymbol{T}) = \lambda(\boldsymbol{AT}) = \mathbf{0} = \boldsymbol{b}$, we get $\lambda \boldsymbol{T} \in \boldsymbol{D}$.

Third, with respect to the inner operation: $\forall \boldsymbol{T}_1, \boldsymbol{T}_2 \in \boldsymbol{D}$, then $\boldsymbol{AT}_1 = \boldsymbol{b} = \mathbf{0}$, $\boldsymbol{AT}_2 = \boldsymbol{b} = \mathbf{0}$, and $\boldsymbol{A}(\boldsymbol{T}_1+\boldsymbol{T}_2) = \boldsymbol{AT}_1 + \boldsymbol{AT}_2 = \mathbf{0} = \boldsymbol{b}$, we get $\boldsymbol{T}_1+\boldsymbol{T}_2 \in \boldsymbol{D}$.

Thus, it also satisfies axiom of closure. In sum, all the subspace axioms are satisfied, and we have proved that $\boldsymbol{D}$ is a subspace of $\mathbb{R}^n$ if and only if $\boldsymbol{b} = \mathbf{0}$.

Exercise 4

(a) Since T is a linear transformation, we have $T(\mathbf{0}+\mathbf{0}) = T(\mathbf{0}) + T(\mathbf{0}) = 2T(\mathbf{0})$, and $T(\mathbf{0}+\mathbf{0}) = T(\mathbf{0})$, we derive $T(\mathbf{0}) = 2T(\mathbf{0})$, thus $T(\mathbf{0}) = \mathbf{0}$.

(b) ***Proof:*** We will prove by induction that, for all $n \in \mathbb{Z}_+$,

$$T(c_1\boldsymbol{v}_1 + ... + c_n\boldsymbol{v}_n) = c_1T(\boldsymbol{v}_1) + ... + c_nT(\boldsymbol{v}_n) \tag{1}$$

Base case: When $n = 1$, $LFS = T(c_1\boldsymbol{v}_1) = T(c_1\boldsymbol{v}_1 + \mathbf{0}) = c_1T(\boldsymbol{v}_1) + T(\mathbf{0})$, and we have proved that $T(\mathbf{0}) = \mathbf{0}$ in (a), thus $LFS = c_1T(\boldsymbol{v}_1)$, and $RHS = c_1T(\boldsymbol{v}_1)$, so both sides are equal and (1) is true for $n = 1$.

Induction step: Let $k \in \mathbb{Z}_+$ be given and suppose (1) is true for $n = k$. Then

$$\begin{aligned}
T(c_1\boldsymbol{v}_1 + ... + c_k\boldsymbol{v}_k + c_{k+1}\boldsymbol{v}_{k+1}) &= T((c_1\boldsymbol{v}_1 + ... + c_k\boldsymbol{v}_k) + c_{k+1}\boldsymbol{v}_{k+1}) \\
&= T(c_1\boldsymbol{v}_1 + ... + c_k\boldsymbol{v}_k) + T(c_{k+1}\boldsymbol{v}_{k+1}) \\
&= c_1T(\boldsymbol{v}_1) + ... + c_kT(\boldsymbol{v}_k) + T(c_{k+1}\boldsymbol{v}_{k+1}) \\
&= c_1T(\boldsymbol{v}_1) + ... + c_kT(\boldsymbol{v}_k) + c_{k+1}T(\boldsymbol{v}_{k+1}).
\end{aligned}$$

Thus, (1) holds for $n = k + 1$, and the proof of the induction step is complete.

Conclusion: By the principle of induction, (1) is true for all $n \in \mathbb{Z}_+$.

(c) As $\{\boldsymbol{v}_1, ..., \boldsymbol{v}_n\}$ is a set of linearly dependent vectors in V, there exist $\lambda_1, ..., \lambda_2$, with at least one $\lambda_n \neq 0$, such that

$$\lambda_1\boldsymbol{v}_1 + \lambda_2\boldsymbol{v}_2 + ... + \lambda_n\boldsymbol{v}_n = \mathbf{0}$$

We know $T : V \to W$ is a linear transformation, $\{\boldsymbol{v}_1, ..., \boldsymbol{v}_n\}$ is a set of vectors in V, and $\boldsymbol{w}_n := T(\boldsymbol{v}_n)$, then $\{\boldsymbol{w}_1, ..., \boldsymbol{w}_n\}$ is a set of vectors in W.

Multiply $\lambda_1\boldsymbol{v}_1 + \lambda_2\boldsymbol{v}_2 + ... + \lambda_n\boldsymbol{v}_n$ by T, we get

$$\begin{aligned}
T(\lambda_1\boldsymbol{v}_1 + \lambda_2\boldsymbol{v}_2 + ... + \lambda_n\boldsymbol{v}_n) &= \lambda_1T(\boldsymbol{v}_1) + \lambda_2T(\boldsymbol{v}_2) + ... + \lambda_nT(\boldsymbol{v}_n) \\
&= \lambda_1\boldsymbol{w}_1 + \lambda_2\boldsymbol{w}_2 + ... + \lambda_n\boldsymbol{w}_n
\end{aligned}$$

$T(\lambda_1\boldsymbol{v}_1 + \lambda_2\boldsymbol{v}_2 + ... + \lambda_n\boldsymbol{v}_n) = T(\mathbf{0})$, and we have proved that $T(\mathbf{0}) = \mathbf{0}$ in (a), then there exist $\lambda_1, ..., \lambda_2$, with at least one $\lambda_n \neq 0$, such that

$$\lambda_1\boldsymbol{w}_1 + \lambda_2\boldsymbol{w}_2 + ... + \lambda_n\boldsymbol{w}_n = \mathbf{0}$$

Thus, $\{\boldsymbol{w}_1, ..., \boldsymbol{w}_n\}$ is a set of linearly dependent vectors in W.

Exercise 5

(a) Let V be the vector space and $\langle\cdot,\cdot\rangle = \Omega : V \times V \to \mathbb{R}$ is symmetric and linear in the first argument, for all $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in V, \lambda, \psi \in \mathbb{R}$ that

$$\begin{aligned}
\Omega(\boldsymbol{x}, \lambda\boldsymbol{y} + \psi\boldsymbol{z}) &= \Omega(\lambda\boldsymbol{y} + \psi\boldsymbol{z}, \boldsymbol{x}) \cdots\cdots\text{(symmetric)} \\
&= \lambda\Omega(\boldsymbol{y}, \boldsymbol{x}) + \psi\Omega(\boldsymbol{z}, \boldsymbol{x}) \cdots\cdots\text{(linear in the first argument)} \\
&= \lambda\Omega(\boldsymbol{x}, \boldsymbol{y}) + \psi\Omega(\boldsymbol{x}, \boldsymbol{z}) \cdots\cdots\text{(symmetric)}
\end{aligned}$$

We get $\Omega(\boldsymbol{x}, \lambda\boldsymbol{y} + \psi\boldsymbol{z}) = \lambda\Omega(\boldsymbol{x}, \boldsymbol{y}) + \psi\Omega(\boldsymbol{x}, \boldsymbol{z})$, then Ω is linear in the second argument as well. Hence, $\langle\cdot,\cdot\rangle$ is a bilinear mapping.

(b) $\langle\cdot,\cdot\rangle$ is **symmetric**: Let $\boldsymbol{x} := [x_1, x_2]^\top, \boldsymbol{y} := [y_1, y_2]^\top \in \mathbb{R}^2$. Then

$$\begin{aligned}\langle\boldsymbol{x},\boldsymbol{y}\rangle &= x_1y_1 + x_2y_2 + 2(x_1y_2 + x_2y_1)\\ &= y_1x_1 + y_2x_2 + 2(y_2x_1 + y_1x_2)\\ &= y_1x_1 + y_2x_2 + 2(y_1x_2 + y_2x_1) = \langle\boldsymbol{y},\boldsymbol{x}\rangle\end{aligned}$$

where we exploited the commutative of addition and multiplication in $\mathbb{R}$. Therefore, $\langle\cdot,\cdot\rangle$ is symmetric.

$\langle\cdot,\cdot\rangle$ is **NOT positive definite**: Consider $\boldsymbol{x} := [1, -1]^\top \in \mathbb{R}^2$. Then

$$\begin{aligned}\langle\boldsymbol{x},\boldsymbol{x}\rangle &= x_1^2 + x_2^2 + 2(x_1x_2 + x_2x_1)\\ &= 1^2 + (-1)^2 + 2\cdot(1\cdot(-1) + (-1)\cdot 1)\\ &= -2\end{aligned}$$

we get $\exists\boldsymbol{x} \in V\backslash\{\boldsymbol{0}\} : \langle\boldsymbol{x},\boldsymbol{x}\rangle < 0$. Hence, $\langle\cdot,\cdot\rangle$ is not positive definite.

$\langle\cdot,\cdot\rangle$ is **bilinear**: Let $\boldsymbol{z} := [z_1, z_2]^\top \in \mathbb{R}^2$ and $\lambda \in \mathbb{R}$. Then

$$\begin{aligned}\langle\boldsymbol{x}+\boldsymbol{y},\boldsymbol{z}\rangle &= (x_1+y_1)z_1 + (x_2+y_2)z_2 + 2((x_1+y_1)z_2 + (x_2+y_2)z_1)\\ &= x_1z_1 + x_2z_2 + 2(x_1z_2 + x_2z_1) + y_1z_1 + y_2z_2 + 2(y_1z_2 + y_2z_1)\\ &= \langle\boldsymbol{x},\boldsymbol{z}\rangle + \langle\boldsymbol{y},\boldsymbol{z}\rangle\end{aligned}$$

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

Thus, $\langle\cdot,\cdot\rangle$ is linear in its first argument. As we have proved in (a), by symmetry, $\langle\cdot,\cdot\rangle$ is bilinear.

Overall, $\langle\cdot,\cdot\rangle$ is **symmetric** and **bilinear**.

Exercise 6

(a) ***Theorem:*** If $\boldsymbol{x}$ and $\boldsymbol{y}$ are orthogonal, then they are linearly independent. (TRUE)

Proof: By contradiction; assume that if $\boldsymbol{x}$ and $\boldsymbol{y}$ are orthogonal, then they are linearly dependent, which means that there exists a non-trivial linear combination, such that $\lambda_1\boldsymbol{x} + \lambda_2\boldsymbol{y} = \boldsymbol{0}$ with at least one $\lambda_n \neq 0$. Since $\boldsymbol{x}$ and $\boldsymbol{y}$ are orthogonal, $\langle\boldsymbol{x},\boldsymbol{y}\rangle = 0$.

Consider the inner product $\langle \lambda_1 \boldsymbol{x} + \lambda_2 \boldsymbol{y}, \boldsymbol{y} \rangle$, apply homogeneity in the first argument, we have

$$\begin{aligned}\langle \lambda_1 \boldsymbol{x} + \lambda_2 \boldsymbol{y}, \boldsymbol{y} \rangle &= \langle \boldsymbol{0}, \boldsymbol{y} \rangle \\ &= \langle 0\boldsymbol{y}, \boldsymbol{y} \rangle \\ &= 0\langle \boldsymbol{y}, \boldsymbol{y} \rangle \\ &= 0\end{aligned}$$

Consider the inner product $\langle \lambda_1 \boldsymbol{x} + \lambda_2 \boldsymbol{y}, \boldsymbol{y} \rangle$, apply linear mapping in the first argument, we have

$$\begin{aligned}\langle \lambda_1 \boldsymbol{x} + \lambda_2 \boldsymbol{y}, \boldsymbol{y} \rangle &= \lambda_1 \langle \boldsymbol{x}, \boldsymbol{y} \rangle + \lambda_2 \langle \boldsymbol{y}, \boldsymbol{y} \rangle \\ &= 0 + \lambda_2 \langle \boldsymbol{y}, \boldsymbol{y} \rangle \\ &= \lambda_2 \langle \boldsymbol{y}, \boldsymbol{y} \rangle\end{aligned}$$

Thus, we have $\lambda_2 \langle \boldsymbol{y}, \boldsymbol{y} \rangle = 0$. Because inner product is positive definite, and $\boldsymbol{y}$ is a non-zero vector, we have $\langle \boldsymbol{y}, \boldsymbol{y} \rangle > 0$, then $\lambda_2 = 0$.

Similarly, consider the inner product $\langle \boldsymbol{x}, \lambda_1 \boldsymbol{x} + \lambda_2 \boldsymbol{y} \rangle$, apply homogeneity in the first argument and linear mapping in the first argument, we have

$$\langle \boldsymbol{x}, \lambda_1 \boldsymbol{x} + \lambda_2 \boldsymbol{y} \rangle = 0 = \lambda_1 \langle \boldsymbol{x}, \boldsymbol{x} \rangle$$

Thus, we have $\lambda_1 \langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0$. Because inner product is positive definite, and $\boldsymbol{x}$ is a non-zero vector, we have $\langle \boldsymbol{x}, \boldsymbol{x} \rangle > 0$, then $\lambda_1 = 0$.

From above, we have $\lambda_1 = \lambda_2 = 0$. However, this contradicts our earlier assumption that at least one $\lambda_n \neq 0$, so our assumption must have been wrong. Therefore, if $\boldsymbol{x}$ and $\boldsymbol{y}$ are orthogonal, then they are linearly independent.

(b) ***Theorem:*** If $\boldsymbol{x}$ and $\boldsymbol{y}$ are linearly independent, then they are orthogonal. (FALSE)
Disproof: Let $\boldsymbol{x'}$, $\boldsymbol{y'}$ be **non-zero** and **orthogonal** vectors in V, and $\boldsymbol{x'} \neq \boldsymbol{y'}$, then we have $\langle \boldsymbol{x'}, \boldsymbol{x'} \rangle > 0$, $\langle \boldsymbol{y'}, \boldsymbol{y'} \rangle > 0$, and $\langle \boldsymbol{x'}, \boldsymbol{y'} \rangle = 0$. From the conclusion in (a), we derive $\boldsymbol{x'}$ and $\boldsymbol{y'}$ are **linearly independent**.

Let $\boldsymbol{x} := \boldsymbol{x'} - \boldsymbol{y'}$ and $\boldsymbol{y} := \boldsymbol{y'}$, I will first prove that $\boldsymbol{x}$ and $\boldsymbol{y}$ are linearly independent.

Proof by contradiction: Assume that $\boldsymbol{x}$ and $\boldsymbol{y}$ are linearly dependent, which means that there exists a non-trivial linear combination, such that $\lambda_1 \boldsymbol{x} + \lambda_2 \boldsymbol{y} = \boldsymbol{0}$ with at least one $\lambda_n \neq 0$. Then

$$\begin{aligned}\lambda_1 \boldsymbol{x} + \lambda_2 \boldsymbol{y} &= \lambda_1 (\boldsymbol{x'} - \boldsymbol{y'}) + \lambda_2 \boldsymbol{y'} \\ &= \lambda_1 \boldsymbol{x'} + (\lambda_2 - \lambda_1) \boldsymbol{y'} = \boldsymbol{0}\end{aligned}$$

We get $\lambda_1 \boldsymbol{x'} + (\lambda_2 - \lambda_1)\boldsymbol{y'} = \mathbf{0}$. Since $\boldsymbol{x'}$ and $\boldsymbol{y'}$ are linearly independent, only the trivial solution exists, thus $\lambda_1 = (\lambda_2 - \lambda_1) = 0$. Then, $\lambda_1 = \lambda_2 = 0$. However, this contradicts our earlier assumption that at least one $\lambda_n \neq 0$, so our assumption must have been wrong. Therefore, $\boldsymbol{x}$ and $\boldsymbol{y}$ are linearly independent.

For the next, I will prove that $\boldsymbol{x}$ and $\boldsymbol{y}$ are not orthogonal.
Consider the inner product $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$, apply linear mapping in the first argument, we have

$$\begin{aligned}\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{x'} - \boldsymbol{y'}, \boldsymbol{y'} \rangle &= \langle \boldsymbol{x'}, \boldsymbol{y'} \rangle + (-1)\langle \boldsymbol{y'}, \boldsymbol{y'} \rangle \\ &= 0 + (-1)\langle \boldsymbol{y'}, \boldsymbol{y'} \rangle \\ &= (-1)\langle \boldsymbol{y'}, \boldsymbol{y'} \rangle\end{aligned}$$

Since $\langle \boldsymbol{y'}, \boldsymbol{y'} \rangle > 0$, we derive $\langle \boldsymbol{x}, \boldsymbol{y} \rangle \neq 0$, then $\boldsymbol{x}$ and $\boldsymbol{y}$ are NOT orthogonal. We have found a counter-example to the original theorem.

Overall, if $\boldsymbol{x}$ and $\boldsymbol{y}$ are linearly independent, we cannot derive that they are orthogonal.

Exercise 7

(a) Yes, $\|\cdot\|_a \overset{\varepsilon}{\sim} \|\cdot\|_a$ for all $\varepsilon \in (0, 1]$.

Since $0 < \varepsilon \leq 1$, then $1 - \varepsilon \geq 0$, and notice $\|\boldsymbol{v}\|_a \geq 0$, we have

$$(1 - \varepsilon)\|\boldsymbol{v}\|_a \geq 0$$

$$\|\boldsymbol{v}\|_a - \varepsilon\|\boldsymbol{v}\|_a \geq 0$$
$$\|\boldsymbol{v}\|_a \geq \varepsilon\|\boldsymbol{v}\|_a \tag{1}$$

Since $0 < \varepsilon \leq 1$, then $\frac{1}{\varepsilon} \geq 1$, and notice $\|\boldsymbol{v}\|_a \geq 0$, multiple (1) by $\frac{1}{\varepsilon}$, we have

$$\frac{1}{\varepsilon}\|\boldsymbol{v}\|_a \geq \|\boldsymbol{v}\|_a \tag{2}$$

Combine (1) and (2), we have : $\varepsilon\|\boldsymbol{v}\|_a \leq \|\boldsymbol{v}\|_a \leq \frac{1}{\varepsilon}\|\boldsymbol{v}\|_a$.

Hence, it is true that $\|\cdot\|_a \overset{\varepsilon}{\sim} \|\cdot\|_a$, *$\varepsilon$-equivalence* is reflexive for all $\varepsilon \in (0, 1]$.

(b) Yes, $\|\cdot\|_a \overset{\varepsilon}{\sim} \|\cdot\|_b$ implies $\|\cdot\|_b \overset{\varepsilon}{\sim} \|\cdot\|_a$ for all $\varepsilon \in (0, 1]$.

From $\|\cdot\|_a \overset{\varepsilon}{\sim} \|\cdot\|_b$, we have $\varepsilon\|\boldsymbol{v}\|_a \leq \|\boldsymbol{v}\|_b \leq \frac{1}{\varepsilon}\|\boldsymbol{v}\|_a$, then

$$\|\boldsymbol{v}\|_b \leq \frac{1}{\varepsilon}\|\boldsymbol{v}\|_a \tag{1}$$

$$\varepsilon\|\boldsymbol{v}\|_a \leq \|\boldsymbol{v}\|_b \tag{2}$$

Since $0 < \varepsilon \leq 1$, and notice $\|\boldsymbol{v}\|_a, \|\boldsymbol{v}\|_b \geq 0$, multiple (1) by ε, we have

$$\varepsilon\|\boldsymbol{v}\|_b \leq \|\boldsymbol{v}\|_a \tag{3}$$

Since $0 < \varepsilon \leq 1$, then $\frac{1}{\varepsilon} \geq 1$, and notice $\|\boldsymbol{v}\|_a, \|\boldsymbol{v}\|_b \geq 0$, multiple (2) by $\frac{1}{\varepsilon}$, we have

$$\|\boldsymbol{v}\|_a \leq \frac{1}{\varepsilon}\|\boldsymbol{v}\|_b \tag{4}$$

Combine (3) and (4), we have : $\varepsilon\|\boldsymbol{v}\|_b \leq \|\boldsymbol{v}\|_a \leq \frac{1}{\varepsilon}\|\boldsymbol{v}\|_b$. Hence, $\|\cdot\|_b \overset{\varepsilon}{\sim} \|\cdot\|_a$.

Overall, $\|\cdot\|_a \overset{\varepsilon}{\sim} \|\cdot\|_b$ implies $\|\cdot\|_b \overset{\varepsilon}{\sim} \|\cdot\|_a$, ε-*equivalence* is symmetric for all $\varepsilon \in (0, 1]$.

(c) ***Proof:*** Since $V = \mathbb{R}^2$, let $\|\boldsymbol{v}\|_1 := |v_1| + |v_2|$ and $\|\boldsymbol{v}\|_2 := \sqrt{v_1^2 + v_2^2}$.

Helper inequation (1): $(|v_1| - |v_2|)^2 \geq 0$, then

$$\begin{aligned}
(|v_1| - |v_2|)^2 &= v_1^2 + v_2^2 - 2|v_1 v_2| \\
&= 2(v_1^2 + v_2^2) - (v_1^2 + v_2^2 + 2|v_1 v_2|) \\
&= 2(\sqrt{v_1^2 + v_2^2})^2 - (|v_1| + |v_2|)^2 \\
&= 2\|\boldsymbol{v}\|_2^2 - \|\boldsymbol{v}\|_1^2 \geq 0
\end{aligned}$$

Since $2\|\boldsymbol{v}\|_2^2 - \|\boldsymbol{v}\|_1^2 \geq 0$, we have $\|\boldsymbol{v}\|_2^2 \geq \frac{1}{2}\|\boldsymbol{v}\|_1^2$, and notice $\|\boldsymbol{v}\|_1, \|\boldsymbol{v}\|_2 \geq 0$, we derive by square root that : $\|\boldsymbol{v}\|_2 \geq \frac{1}{\sqrt{2}}\|\boldsymbol{v}\|_1$, then $\frac{\|\boldsymbol{v}\|_2}{\|\boldsymbol{v}\|_1} \geq \frac{1}{\sqrt{2}}$.

Consider $\varepsilon \in (0, 1]$ that makes $\varepsilon\|\boldsymbol{v}\|_1 \leq \|\boldsymbol{v}\|_2$, then $\varepsilon \leq \frac{\|\boldsymbol{v}\|_2}{\|\boldsymbol{v}\|_1}$. Since $\frac{\|\boldsymbol{v}\|_2}{\|\boldsymbol{v}\|_1} \geq \frac{1}{\sqrt{2}}$, the largest ε possible is : $\varepsilon = \frac{1}{\sqrt{2}}$, and it makes

$$\frac{1}{\sqrt{2}}\|\boldsymbol{v}\|_1 \leq \|\boldsymbol{v}\|_2 \tag{1}$$

Helper inequation (2): $v_1^2 + v_2^2 + 4|v_1 v_2| \geq 0$, then

$$2v_1^2 + 2v_2^2 + 4|v_1 v_2| \geq v_1^2 + v_2^2$$

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

Notice $\|\boldsymbol{v}\|_1, \|\boldsymbol{v}\|_2 \geq 0$, we derive by square root that

$$\sqrt{2}\|\boldsymbol{v}\|_1 \geq \|\boldsymbol{v}\|_2 \tag{2}$$

Combine (1) and (2), we have : $\frac{1}{\sqrt{2}}\|\boldsymbol{v}\|_1 \leq \|\boldsymbol{v}\|_2 \leq \sqrt{2}\|\boldsymbol{v}\|_1$.

Hence, for $V = \mathbb{R}^2$, it is true that $\|\cdot\|_1 \overset{\varepsilon}{\sim} \|\cdot\|_2$ for the the largest possible $\varepsilon = \frac{1}{\sqrt{2}}$.

Exercise 8

(a) We perform Gaussian Elimination on the augmented matrix $\left[\begin{array}{cc|c} 1 & 2 & 12 \\ 1 & 1 & 12 \\ 1 & 0 & 18 \end{array}\right]$ to get the row-echelon form.

$$\begin{bmatrix} 1 & 2 & 12 \\ 1 & 1 & 12 \\ 1 & 0 & 18 \end{bmatrix} \xrightarrow[R_3=R_1-R_3]{R_2=R_1-R_2} \begin{bmatrix} 1 & 2 & 12 \\ 0 & 1 & 0 \\ 0 & 2 & -6 \end{bmatrix} \xrightarrow{R_3=2R_2-R_3} \begin{bmatrix} 1 & 2 & 12 \\ 0 & 1 & 0 \\ 0 & 0 & 6 \end{bmatrix}$$

We can get the REF $\left[\begin{array}{cc|c} 1 & 2 & 12 \\ 0 & 1 & 0 \\ 0 & 0 & 6 \end{array}\right]$. From the third row we have $0+0=6$, which is a contradiction, hence no solution exists. Thus, $\begin{bmatrix} 12 \\ 12 \\ 18 \end{bmatrix}$ cannot be represented by a linear combination of $\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} \right\}$, then $\boldsymbol{x} \notin U$.

(b) First, we determine a basis of U. Writing the spanning vectors as the columns of a matrix A, we use Gaussian elimination to bring A into reduced row echelon form:

$$\begin{bmatrix} 1 & 2 \\ 1 & 1 \\ 1 & 0 \end{bmatrix} \xrightarrow[R_3=R_1-R_3]{R_2=R_1-R_2} \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 0 & 2 \end{bmatrix} \xrightarrow[R_3=R_3-2R_2]{R_1=R_1-R_3} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

We get the RREF: $\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$. From here, we see that the both two columns are pivot columns, i.e., these two vectors in the generating set of U form a basis of U:

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

Now, we define

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

where we define two basis vectors $\boldsymbol{b}_1, \boldsymbol{b}_2$ of U as the columns of $\boldsymbol{B}$.

We know that the projection of $\boldsymbol{x}$ on U exists and we define $\boldsymbol{p} := \pi_U(\boldsymbol{x})$. Moreover, we know that $\boldsymbol{p} \in U$. We define $\boldsymbol{\lambda} := [\lambda_1, \lambda_2]^\top \in \mathbb{R}^2$, such that $\boldsymbol{p}$ can be written $\boldsymbol{p} = \lambda_1 \boldsymbol{b}_1 + \lambda_2 \boldsymbol{b}_2 = \boldsymbol{B}\boldsymbol{\lambda}$.

As $\boldsymbol{p}$ is the orthogonal projection of $\boldsymbol{x}$ onto U, then $\boldsymbol{x} - \boldsymbol{p}$ is orthogonal to all the basis vectors of U, so that

$$\boldsymbol{B}^\top(\boldsymbol{x} - \boldsymbol{B}\boldsymbol{\lambda}) = \boldsymbol{0}$$

Therefore,

$$\boldsymbol{B}^\top \boldsymbol{B}\boldsymbol{\lambda} = \boldsymbol{B}^\top \boldsymbol{x}$$

Solving in $\boldsymbol{\lambda}$ the inhomogeneous system $\boldsymbol{B}^\top \boldsymbol{B}\boldsymbol{\lambda} = \boldsymbol{B}^\top \boldsymbol{x}$ gives us a single solution

$$\begin{aligned}
\boldsymbol{\lambda} = (\boldsymbol{B}^\top \boldsymbol{B})^{-1} \boldsymbol{B}^\top \boldsymbol{x} &= (\begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ 1 & 1 \\ 1 & 0 \end{bmatrix})^{-1} \cdot \begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 12 \\ 12 \\ 18 \end{bmatrix} \\
&= \begin{bmatrix} \frac{5}{6} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 12 \\ 12 \\ 18 \end{bmatrix} \\
&= \begin{bmatrix} -\frac{1}{6} & \frac{1}{3} & \frac{5}{6} \\ \frac{1}{2} & 0 & -\frac{1}{2} \end{bmatrix} \cdot \begin{bmatrix} 12 \\ 12 \\ 18 \end{bmatrix} \\
&= \begin{bmatrix} 17 \\ -3 \end{bmatrix}
\end{aligned}$$

and, therefore, the desired projection

$$\boldsymbol{p} = \boldsymbol{B}\boldsymbol{\lambda} = \begin{bmatrix} 1 & 2 \\ 1 & 1 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 17 \\ -3 \end{bmatrix} = \begin{bmatrix} 11 \\ 14 \\ 17 \end{bmatrix} \in U$$

Overall, the orthogonal projection is $\pi_U(\boldsymbol{x}) = \begin{bmatrix} 11 \\ 14 \\ 17 \end{bmatrix}$.

(c) Suppose that for $c_1, c_2 \in \mathbb{R}$, such that

$$\pi_U(\boldsymbol{x}) = \begin{bmatrix} 11 \\ 14 \\ 17 \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$$

This can be turned into a system of equations shown below:

$$\begin{aligned} c_1 + 2c_2 &= 11 \\ c_1 + c_2 &= 14 \\ c_1 &= 17 \end{aligned}$$

Solving this we get $c_1 = 17$ and $c_2 = -3$.

Therefore, $\pi_U(\boldsymbol{x})$ can be written as a linear combination as

$$\pi_U(\boldsymbol{x}) = 17\left[1, 1, 1\right]^\top - 3\left[2, 1, 0\right]^\top$$

(d) The minimum distance is simply the length of $\boldsymbol{x} - \boldsymbol{p}$:

$$\begin{aligned} \|\boldsymbol{x} - \boldsymbol{p}\|_2 &= \left\| \begin{bmatrix} 12 \\ 12 \\ 18 \end{bmatrix} - \begin{bmatrix} 11 \\ 14 \\ 17 \end{bmatrix} \right\|_2 \\ &= \left\| \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \right\|_2 \\ &= \sqrt{1^2 + (-2)^2 + 1^2} \\ &= \sqrt{6} \end{aligned}$$

Overall, the minimum distance is $d(\boldsymbol{x}, U) = \sqrt{6}$.