The Australian National University
School of Computing

Semester 2, 2023
Theory Assignment 1 of 4

COMP3670/6670: Introduction to Machine Learning

Release Date. Aug 3th, 2023
Due Date. 11:59pm, Aug 28th, 2023
Maximum credit. 100

Exercise 1 Properties of Matrices (2+2+2+3+3+4+3 credits)

(a) Let $\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}$ be a square matrix. Show that $\mathbf{A}$ is symmetric.

Solution. $\mathbf{A}^T = \mathbf{A}$, so it is symmetric.

(b) Compute the square of $\mathbf{A}$, that is $\mathbf{A}^2$ and show that $\mathbf{A}^2$ is also symmetric.

Solution. $\mathbf{A}^2 = \begin{bmatrix} 14 & 25 & 31 \\ 25 & 45 & 56 \\ 31 & 56 & 70 \end{bmatrix}$. ${\mathbf{A}^2}^T = \mathbf{A}^2$ so it is symmetric.

(c) Is it true for any symmetric matrix $\mathbf{A}$, $\mathbf{A}^2$ is also symmetric? Show your working.

Solution. True. As $\left(\mathbf{A}^2\right)^T = (\mathbf{A} \cdot \mathbf{A})^T = \mathbf{A}^T \cdot \mathbf{A}^T = \mathbf{A} \cdot \mathbf{A} = \mathbf{A}^2$.

(d) Let $\mathbf{A}$ be a square matrix and $f(\mathbf{X})$ and $g(\mathbf{X})$ be n-th order polynomials, defined by $\sum_{i=0}^{n} a_i \mathbf{X}^i$ where a_i are arbitrary real numbers. Show that the matrices $f(\mathbf{A})$ and $g(\mathbf{A})$ commute, i.e, $f(\mathbf{A})g(\mathbf{A}) = g(\mathbf{A})f(\mathbf{A})$ for arbitrary order n.

Solution. Let $f(\mathbf{A}) = \sum_{i=0}^{n} a_i \mathbf{A}^i$, $g(\mathbf{A}) = \sum_{j=0}^{n} b_j \mathbf{A}^j$. $f(\mathbf{A})g(\mathbf{A}) = \sum_{i=0}^{n} a_i \mathbf{A}^i \sum_{j=0}^{n} b_j \mathbf{A}^j = \sum_{i=0}^{n} \sum_{j=0}^{n} a_i \mathbf{A}^i b_j \mathbf{A}^j = \sum_{j=0}^{n} \sum_{i=0}^{n} b_j \mathbf{A}^j a_i \mathbf{A}^i = \sum_{j=0}^{n} b_j \mathbf{A}^j \sum_{i=0}^{n} a_i \mathbf{A}^i = g(\mathbf{A})f(\mathbf{A})$.

(e) Let $\mathbf{A}$ and $\mathbf{B}$ be rectangular matrices of orders $n \times k$ and $r \times s$, respectively. The matrix of order $nr \times ks$ represented in a block form as

$$\begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \dots & a_{1k}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \dots & a_{2k}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}\mathbf{B} & a_{n2}\mathbf{B} & \dots & a_{nk}\mathbf{B} \end{bmatrix}$$

is called the Kronecker product $\mathbf{A} \otimes \mathbf{B}$ of the matrices $\mathbf{A}$ and $\mathbf{B}$.

Let

$$\mathbf{X} = \begin{bmatrix} 8 & 1 & 6 \\ 3 & 5 & 7 \\ 4 & 9 & 2 \end{bmatrix}.$$

$\mathbf{X}$ is a so-called magic square, since its row sums, column sums, principal diagonal sum, and principal counter diagonal sum are all equal. Is $\mathbf{X} \otimes \mathbf{X}$ a magic square?

Solution. Yes. Computation should be shown.

$$\mathbf{X} \otimes \mathbf{X} = \begin{bmatrix} 64 & 8 & 48 & 8 & 1 & 6 & 48 & 6 & 36 \\ 24 & 40 & 56 & 3 & 5 & 7 & 18 & 30 & 42 \\ 32 & 72 & 16 & 4 & 9 & 2 & 24 & 54 & 12 \\ 24 & 3 & 18 & 40 & 5 & 30 & 56 & 7 & 42 \\ 9 & 15 & 21 & 15 & 25 & 35 & 21 & 35 & 49 \\ 12 & 27 & 6 & 20 & 45 & 10 & 28 & 63 & 14 \\ 32 & 4 & 24 & 72 & 9 & 54 & 16 & 2 & 12 \\ 12 & 20 & 28 & 27 & 45 & 63 & 6 & 10 & 14 \\ 16 & 36 & 8 & 36 & 81 & 18 & 8 & 18 & 4 \end{bmatrix}.$$

The row, column and diagonal sums are equal.

(f) Determine if $\mathbf{X} \otimes \mathbf{X}$ is a magic square for any magic matrix $\mathbf{X}$ of order $n \times n$. Show an example of a magic square with n=2 for which $\mathbf{X} \otimes \mathbf{X}$ is also a magic square.

Solution. Consider a general case where $\mathbf{X} \in \mathbb{R}^{n \times n}$. $\mathbf{X} \otimes \mathbf{X} = \begin{bmatrix} X_{11}\mathbf{X} & X_{12}\mathbf{X} & \dots & X_{1n}\mathbf{X} \\ X_{21}\mathbf{X} & X_{22}\mathbf{X} & \dots & X_{2n}\mathbf{X} \\ \vdots & \vdots & \ddots & \vdots \\ X_{n1}\mathbf{X} & X_{n2}\mathbf{X} & \dots & X_{nn}\mathbf{X} \end{bmatrix}$.

Let's consider the rows first. An arbitrary row of the result matrix is given as $\begin{bmatrix} X_{i1}\mathbf{X}_t & X_{i2}\mathbf{X}_t & \cdots & X_{in}\mathbf{X}_t \end{bmatrix}$, where $\mathbf{X}_t$ is the t-th row of matrix $\mathbf{X}$. The sum of this row is

$$\sum_{j=1}^{n} X_{ij} \sum_{k=1}^{n} X_{tk}$$

We know $\mathbf{X}$ itself is also a magic matrix. So for arbitrary row t,

$$\sum_{k=1}^{n} X_{tk} = C$$

and for arbitrary column i,

$$\sum_{j=1}^{n} X_{ij} = C$$

where C is a fixed constant. Thus,

$$\sum_{j=1}^{n} X_{ij} \sum_{k=1}^{n} X_{tk} = C \sum_{j=1}^{n} X_{ij} = C^2$$

Similarly, we can show that arbitrary column sum, the diagonal sums are all equal to C^2. So, we have shown that $\mathbf{X} \otimes \mathbf{X}$ is a magic square for all magic square $\mathbf{X}$ of size $n \times n$. Hence, when

$$\mathbf{X} = \begin{bmatrix} 8 & 1 & 6 \\ 3 & 5 & 7 \\ 4 & 9 & 2 \end{bmatrix}$$

$\mathbf{X} \otimes \mathbf{X}$ is a magic square.
When $n = 2$ we don't have a traditional magic square with distinct entries. However, we can always choose

$$\mathbf{Y} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

for which $\mathbf{Y} \otimes \mathbf{Y}$ has the row sums, column sums, principal diagonal sum, and principal counter diagonal sum all equal.

(g) Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$. What conditions should $\mathbf{x}, \mathbf{y}$ satisfy such that $\mathbf{x} \otimes \mathbf{y} \otimes \mathbf{x} = \mathbf{y} \otimes \mathbf{x} \otimes \mathbf{y}$?

Solution. Let $\mathbf{x} = \begin{bmatrix} a & b \end{bmatrix}^T$, $\mathbf{y} = \begin{bmatrix} c & d \end{bmatrix}^T$. $\mathbf{x} \otimes \mathbf{y} \otimes \mathbf{x} = \begin{bmatrix} aca & acb & ada & adb & bca & bcb & bda & bdb \end{bmatrix}^T$. $\mathbf{y} \otimes \mathbf{x} \otimes \mathbf{y} = \begin{bmatrix} cac & cad & cbc & cbd & dac & dad & dbc & dbd \end{bmatrix}^T$. They are equal if the they are equal elementwise. Apparently, if $\mathbf{x} = \mathbf{y}$, the equality holds. Suppose they are not equal, meaning $a \neq c \vee b \neq d$. Suppose $a \neq c \wedge b = d$, according to the elementwise equal property, the first elements in the two results should be equal, namely $cac = aca \implies ac(a - c) = 0$. Given $a \neq c$, we should get $ac = 0$. Also, given $cbd = adb \implies (a - c)bd = 0$, since we have $a \neq c \wedge b = d$, $b = d = 0$. This means if any one of a, c is 0 (but not both), and b, d are 0, the equality holds. Similar reasoning can be applied to prove under the case $a = c \wedge b \neq d$, if any one of b, d is 0 (but not both), and a, c are 0, the equality holds. Now consider the case $a \neq c \wedge b \neq d$. This implies $\mathbf{x} = \mathbf{0}, c, d \neq 0$ or $\mathbf{y} = \mathbf{0}, a, b \neq 0$. In conclusion, $\mathbf{x} \otimes \mathbf{y} \otimes \mathbf{x} = \mathbf{y} \otimes \mathbf{x} \otimes \mathbf{y}$ if and only if $\mathbf{x} = \mathbf{y}$, or exactly one of a, b, c, d is not 0, or $\mathbf{x} = \mathbf{0}, c, d \neq 0$, or $\mathbf{y} = \mathbf{0}, a, b \neq 0$.

Exercise 2 **Solving Linear Systems** (3+3 credits)

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

(a)

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 \\ 6 & 2 & -5 \\ 2 & 1 & -1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 5 \\ -2 \\ 3 \end{bmatrix}$$

Solution. $\mathcal{S} = \left\{ \mathbf{x} = \frac{1}{5} \cdot \begin{bmatrix} 7 \\ 19 \\ 18 \end{bmatrix} \right\}$

(b)

$$\mathbf{A} = \begin{bmatrix} 4 & 3 & 2 & 2 & -2 \\ 0 & 1 & 2 & 2 & 6 \\ 3 & 2 & 1 & 1 & -3 \\ -1 & 0 & 1 & 1 & 1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 5 \\ 23 \\ -2 \\ 16 \end{bmatrix}$$

Solution. $\mathcal{S} = \left\{ \mathbf{x} = \begin{bmatrix} -16 \\ 23 \\ 0 \\ -0 \\ 0 \end{bmatrix} + \lambda_1 \begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \lambda_2 \begin{bmatrix} 1 \\ -2 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \lambda_1, \lambda_2 \in \mathbb{R} \right\}$

Exercise 3 **Inverses and rank** (3+4+6 credits)

(a) Let $\mathbf{A} \in \mathbb{M}_{n\times n}(\mathbb{R})$ be an invertible matrix. Show that the transpose of the inverse of $\mathbf{A}$, denoted $(\mathbf{A}^{-1})^T$, is equal to the inverse of the transpose of $\mathbf{A}$, denoted $(\mathbf{A}^T)^{-1}$.

Solution. Since $(\mathbf{A}^{-1})^T \cdot \mathbf{A}^T = (\mathbf{A} \cdot \mathbf{A}^{-1})^T = \mathbf{I}$, and $\mathbf{A}$ is invertible, $(\mathbf{A}^{-1})^T$ is the inverse of $\mathbf{A}^T$.

(b) Find the values of $[a, b, c]^T \in \mathbb{R}^3$ for which the inverse of the following matrix exists.

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

Solution. Perform Gaussian elimination on $[\mathbf{A}|\mathbf{I}]$ to get $\left[\mathbf{I}|\mathbf{A}^{-1}\right]$. We should obtain an intermediate result

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & a-1 & c-1 & 0 & 1 & -1 \\ 0 & 0 & b-1 & 1 & 0 & -1 \end{bmatrix}$$

To make the left half of the matrix possibly be an identity matrix, $a \neq 1, b \neq 1$. c can be any real number.

(c) Let $\mathbf{A}$ be an arbitrary matrix in $\mathbb{M}_{m\times n}(\mathbb{R})$, where m and n denote the number of rows and columns of $\mathbf{A}$, respectively. Prove that $\mathbf{rk}(\mathbf{A}) = \mathbf{rk}(\mathbf{A}^T)$, where $\mathbf{rk}(\mathbf{A})$ denotes the rank of matrix $\mathbf{A}$.

Solution. Using Elementary Row Operations (EROs) which do not change the row space, we can switch to the Reduced Row Echelon Form (RREF) A' of A. After switching some columns if necessary, we can assume it looks as follows:

$$A' = \left[\begin{array}{c|c} I_r & C_{r,n-r} \\ \hline 0_{m-r,r} & 0_{m-r,n-r} \end{array}\right]$$

for some $r \times (m-r)$ matrix C. Now, the row-rank of A' is r, since the first r rows are linearly independent (the 1's in I_r in the upper-left corner are in different columns). But also, the column-rank of A' is r since the first r columns form the standard basis for $\mathbb{R}^r$, and the columns of C are linear combinations of these. Thus, $\mathrm{rk}(A) = \mathrm{rk}(A^T)$ as they both equal r, completing the proof.
We can accept multiple solutions here, since this isn't the only proof.

Exercise 4 **Subspaces** (2+2+2+3+4 credits)

(a) Which of the following sets are subspaces of $\mathbb{R}^n$? 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.)

(i) $A = \{(x_1, x_2, \dots, x_n) \in \mathbb{R}^n : x_1 \geq 0, x_2 \geq 0, \dots, x_n \geq 0\}$

(ii) $B = \{(x_1, x_2, \dots, x_n) \in \mathbb{R}^n : \text{at least one } x_i \text{ is irrational}\}$

(iii) $C = \left\{(x_1, x_2, \dots, x_n) \in \mathbb{R}^n : \sum_{i=1}^{n} (-1)^{i+1} x_i \geq 0\right\}$

(iv) $D =$ 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. (i) False. We know for an arbitrary vector space V, if $v \in V \implies cv \in V$ for $c \in R$. We can see for $c = -1$ we do not have $v \in A \implies -1v \in A$ since if $x_1 \geq x_2$ then $-1x_1 \leq -1x_2$

Solution. False. We can see $v = [\pi, \pi]^T \in B$ so for all $c \in \mathbb{R}$ we must have $cv \in B$. We know $c = \frac{1}{\pi} \in \mathbb{R}$ but $cv = \frac{1}{\pi}[\pi, \pi]^T = [1, 1]^T \notin B$

Solution. False, We have $[2, 1] \in C$ since $(-1)^2 * 2 + -1 * 1 = 2 - 1 = 1 \geq 0$. However, $-1[2, 1] = [-2, -1] \in C$ since $(-1)^2 * -2 + -1 * -1 = -2 + 1 = -1 < 0$.

Solution. We have the D is a vector space if and only if $b = 0$ (The zero vector).

If $b \neq 0$ then $0 \notin D$ since $A0 = 0 \neq b$ so D cannot be a subspace of $\mathbb{R}^n$.

For the other direction we must show that $x \in D \implies cx \in D \forall c \in \mathbb{R}$ and $x, y \in D \implies x + y \in D$. Both come from the linearity of the operator $\mathbf{A}$.

We see if $x \in D$ then $Ax = b = 0$ so $A(cx) = cAx = c0 = 0 = b$ so $cx \in D$.

Likewise if $x, y \in D$ then $Ax = 0$ and $Ay = 0$ so $A(x + y) = Ax + Ay = 0 + 0 = 0 = b$ so $x + y \in D$.

We satisfied the conditions so D is a subspace when $b = 0$.

Thus D is a subspace iff $b = 0$.

(b) Let V be an inner product space, and let W be a subspace of V. The orthogonal complement of W, denoted $W^\perp$, is defined as the set of all vectors in V that are orthogonal to every vector in W. Show that $W^\perp$ is also a subspace of V.

Solution. We need to show that $v \in W^\perp \implies cv \in W^\perp$ for all $c \in \mathbb{R}$ and that $u, v \in W^\perp \implies u + v \in W^\perp$. We know $x \in W^\perp :=< x, w >= 0$ for all $w \in W$. Using this we see that if $v \in W^\perp$ then $< v, w >= 0$ for all $w \in W$ so $< cv, w >= c < v, w >= c0 = 0$ for all $w \in W$ so $cv \in W^\perp$. Likewise we get if $u, v \in W^\perp$ then both $< u, w >= 0$ for all $w \in W$ and $< v, w >= 0$ for all $w \in W$ so $< u + v, w >=< u, w > + < v, w >= 0 + 0 = 0$ for all $w \in W$ so $u + v \in W^\perp$. Thus we have satisfied the two conditions for $W^\perp$ to be a subspace.

Exercise 5 **Linear Independence** (4+4+4+4+4 credits)

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

The *image* of T is defined as:

$$\mathbf{Im}(T) = \{w \in W \mid \exists v \in V \text{ such that } w = T(v)\}.$$

The *kernel* of T is defined as:

$$\mathbf{Ker}(T) = \{v \in V \mid T(v) = 0\}.$$

We say that T is *injective* if for all $u, v \in V$, $T(u) = T(v)$ implies $u = v$.

(a) Show 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 $n \geq 1$, show 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

$$T(c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n) = c_1T(\mathbf{v}_1) + \ldots + c_nT(\mathbf{v}_n)$$

Solution. We proceed by induction. The base case follows immediately from the definition of linearity of T, as $T(c_1\mathbf{v}_1) = c_1T(\mathbf{v}_1)$. Step case, assume that

$$T(c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n) = c_1T(\mathbf{v}_1) + \ldots + c_nT(\mathbf{v}_n) \qquad \text{(Induction Hypothesis)}$$

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

$$T(c_1\mathbf{v}_1 + \ldots + c_n\mathbf{v}_n + c_{n+1}\mathbf{v}_{n+1}) = c_1T(\mathbf{v}_1) + \ldots + c_nT(\mathbf{v}_n) + c_{n+1}T(\mathbf{v}_{n+1})$$

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) Prove that $\mathbf{Im}(T)$ is a vector subspace of W and $\mathbf{Ker}(T)$ is a vector subspace of V.

Solution.
Ker(T):
We see that if $x \in \mathbf{Ker}(\mathrm{T})$ then $Tx = 0$ then

$$T(cx) = cTx = c0 = 0$$

so $cx \in \mathbf{Ker}(\mathrm{T})$.
Likewise, if $x \in \mathbf{Ker}(\mathrm{T})$ and $y \in \mathbf{Ker}(\mathrm{T})$ then $Tx = 0$ and $Ty = 0$. Thus,

$$T(x+y) = Tx + Ty = 0 + 0 = 0$$

so $x + y \in \mathbf{Ker}(\mathrm{T})$.
Thus $\mathbf{Ker}(\mathrm{T})$ is a subspace of V.

$\mathbf{Im}(T)$:
We know if $v \in \mathbf{Im}(T)$ then $\exists x \in V$ where $Tx = v$. We know that $x \in V \implies cx \in V$ (V is vector space) so we have $cx \in V$ such that

$$T(cx) = cTx = cv$$

Thus $v \in \mathbf{Im}(T) \implies cv \in \mathbf{Im}(T)$.
Likewise if $u \in \mathbf{Im}(T)$ then $\exists y \in V$ where $Ty = u$. We know $x + y \in V$ since V is a vector space. Since,

$$T(x+y) = Tx + Ty = v + u$$

we see $v + u \in \mathbf{Im}(T)$. Thus we have $v \in \mathbf{Im}(T)$ and $u \in \mathbf{Im}(T) \implies v + u \in \mathbf{Im}(T)$

Thus $\mathbf{Im}(T)$ is a subspace.

(d) The Rank-Nullity Theorem states that for a linear map $T : V \to W$, the dimension of the finite-dimensional domain V is equal to the sum of the dimensions of the kernel and the image of T, i.e.,

$$\mathbf{dim}(V) = \mathbf{dim}(\mathbf{Ker}(T)) + \mathbf{dim}(\mathbf{Im}(T)).$$

Give an example of a linear map T such that $\mathbf{dim}(\mathbf{Im}(T)) = 3$ and $\mathbf{dim}(\mathbf{Ker}(T)) = 2$.

Solution. The linear map defined by the matrix,

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

(e) Consider the transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ defined by $T(\mathbf{x}) = \mathbf{A}\mathbf{x}$, where

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

with $[a, b, c]^T \in \mathbb{R}^3$. Find the conditions on a, b, and c for which this transformation is injective.

Solution. To check for injectivity, we can examine the kernel of the transformation T. The transformation T is injective if and only if its kernel contains only the zero vector. The kernel is defined as:

$$\mathbf{Ker}(T) = \{\mathbf{x} \in \mathbb{R}^3 \mid \mathbf{A}\mathbf{x} = \mathbf{0}\}$$

This leads to the system of equations:

$$\begin{aligned} x_1 + ax_2 + bx_3 &= 0, \\ x_1 + x_2 + cx_3 &= 0, \\ x_1 + x_2 + x_3 &= 0. \end{aligned}$$

Simplifying, we find:

$$\begin{aligned} (a-1)x_2 + (b-1)x_3 &= 0, \\ (c-1)x_3 &= 0. \end{aligned}$$

For the transformation to be injective, we need a trivial kernel. Suppose $c = 1$, then x_3 can have multiple values, which makes the kernel non-trivial. So $c \neq 1$. Suppose $a = 1$, then x_2 can have multiple values, which makes the kernel non-trivial. So $a \neq 1$. The value of b does not matter. In conclusion,

$$a \neq 1 \wedge c \neq 1$$

Exercise 6 **Inner Products** (3+3+4+3+5 credits)

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

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

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

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

(b) Given a 2×2 rotation matrix $\mathbf{R}$ represented as

$$\mathbf{R} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix},$$

show that it preserves the standard inner product, i.e., for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$, we have $\mathbf{x}^T\mathbf{y} = (\mathbf{R}\mathbf{x})^T(\mathbf{R}\mathbf{y})$.

Solution. RHS $= \mathbf{x}^T\mathbf{R}^T\mathbf{R}\mathbf{y} = \mathbf{x}^T \begin{bmatrix} \cos^2\theta + \sin^2\theta & 0 \\ 0 & \cos^2\theta + \sin^2\theta \end{bmatrix} \mathbf{y} = \mathbf{x}^T \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \mathbf{y} = \mathbf{x}^T\mathbf{y}$

(c) Now, let us consider an inner product in $\mathbb{R}^2$ defined by the 2×2 matrix

$$\mathbf{D} = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}.$$

Find the matrix $\mathbf{D}'$ (in terms of $\mathbf{R}$ and $\mathbf{D}$) such that the inner product defined by $\mathbf{D}$ is preserved under the rotation by $\mathbf{R}$, i.e., for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$, we have $\mathbf{x}^T\mathbf{D}\mathbf{y} = (\mathbf{R}\mathbf{x})^T\mathbf{D}'(\mathbf{R}\mathbf{y})$.

Solution. $\mathbf{D}' = \mathbf{R}\mathbf{D}\mathbf{R}^T$. Points are given for showing that $(\mathbf{R}\mathbf{x})^T\mathbf{D}'(\mathbf{R}\mathbf{y}) = \mathbf{x}^T(\mathbf{R}^T\mathbf{D}'\mathbf{R})\mathbf{y}$. Since this must be equal to $\mathbf{x}^T\mathbf{D}\mathbf{y}$, it implies that $\mathbf{D} = \mathbf{R}^T\mathbf{D}'\mathbf{R}$. The result follows.

(d) For $\theta = \pi/4$, compute $\mathbf{D}'$ explicitly.

Solution. To find $\mathbf{D}'$, we start with the formula $\mathbf{D}' = \mathbf{R}\mathbf{D}\mathbf{R}^T$. The given rotation matrix $\mathbf{R}$ for $\theta = \frac{\pi}{4}$ is

$$\mathbf{R} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

and the given matrix $\mathbf{D}$ is

$$\mathbf{D} = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$$

First, we calculate $\mathbf{R}\mathbf{D}$:

$$\mathbf{R}\mathbf{D} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} \times 2 + \left(-\frac{1}{\sqrt{2}}\right) \times 1 & \frac{1}{\sqrt{2}} \times 1 + \left(-\frac{1}{\sqrt{2}}\right) \times 3 \\ \frac{1}{\sqrt{2}} \times 2 + \frac{1}{\sqrt{2}} \times 1 & \frac{1}{\sqrt{2}} \times 1 + \frac{1}{\sqrt{2}} \times 3 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{3}{\sqrt{2}} & \frac{3}{\sqrt{2}} \end{bmatrix}$$

Next, we multiply by $\mathbf{R}^T$:

$$\mathbf{D}' = \mathbf{R}\mathbf{D}\mathbf{R}^T$$

$$= \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{3}{\sqrt{2}} & \frac{3}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$= \begin{bmatrix} 1.5 & -0.5 \\ -0.5 & 3.5 \end{bmatrix}$$

This gives us the explicit form of $\mathbf{D}'$ that preserves the inner product under rotation by $\mathbf{R}$.

(e) Consider $\mathbf{u} = [1, 1]^T \in \mathbb{R}^2$ and $\mathbf{v} = [2, -1]^T \in \mathbb{R}^2$. Compute the angle between $\mathbf{u}$ and $\mathbf{v}$ under the inner product defined by $\mathbf{D}$, and the angle between $\mathbf{R}\mathbf{u}$ and $\mathbf{R}\mathbf{v}$ under the inner product defined by $\mathbf{D}'$.

Solution. The angle ω between vectors $\mathbf{u}$ and $\mathbf{v}$ under the inner product defined by $\mathbf{D}$ is computed using:

$$\cos(\omega) = \frac{\mathbf{u}^T\mathbf{D}\mathbf{v}}{\sqrt{\mathbf{u}^T\mathbf{D}\mathbf{u}}\sqrt{\mathbf{v}^T\mathbf{D}\mathbf{v}}} \approx 0.286$$

ω is approximately 1.28 radians or $73.40°$ degrees. Similarly, the angle ω' between $\mathbf{R}\mathbf{u}$ and $\mathbf{R}\mathbf{v}$ under the inner product defined by $\mathbf{D}'$ is:

$$\cos(\omega') = \frac{\mathbf{R}\mathbf{u}^T\mathbf{D}'\mathbf{R}\mathbf{v}}{\sqrt{\mathbf{R}\mathbf{u}^T\mathbf{D}'\mathbf{R}\mathbf{u}}\sqrt{\mathbf{R}\mathbf{v}^T\mathbf{D}'\mathbf{R}\mathbf{v}}} \approx 0.286$$

The angle ω' in radians is also approximately 1.28, or 73.40° in degrees. The angles ω and ω' are the same, consistent with the geometric interpretation that the angle between the vectors is invariant under rotation.

Exercise 7 **Orthogonality** (7+4 credits)

(a) 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.

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) Determine if the 'vectors' defined by the functions $p(x) = 3x^2 - 1$ and $q(x) = 2x + 1$ in the inner product space of continuous functions on the interval $[0, 1]$ with the inner product defined by $\langle f, g \rangle = \int_0^1 f(x)g(x)\,dx$ are orthogonal.

You may find the formulae helpful:

$$\int_a^b \alpha x^n \, dx = \left[\frac{\alpha x^{n+1}}{n+1} \right]_a^b = \frac{\alpha b^{n+1}}{n+1} - \frac{\alpha a^{n+1}}{n+1}$$

$$\int_a^b (f(x) + g(x)) \, dx = \int_a^b f(x) \, dx + \int_a^b g(x) \, dx$$

Solution. To determine if the functions $p(x) = 3x^2 - 1$ and $q(x) = 2x + 1$ are orthogonal, we evaluate their inner product using the formula:

$$\langle p, q \rangle = \int_0^1 (3x^2 - 1)(2x + 1) \, dx$$

We write the product explicitly:

$$(3x^2 - 1)(2x + 1) = 6x^3 + 3x^2 - 2x - 1$$

We integrate this expression from 0 to 1:

$$\int_0^1 (6x^3 + 3x^2 - 2x - 1)\, dx$$

$$= \left[\frac{6x^4}{4} + \frac{3x^3}{3} - \frac{2x^2}{2} - x \right]_0^1$$

$$= \frac{6}{4} + 1 - 1 - 1 = \frac{1}{2}$$

Since the inner product $\langle p, q \rangle$ is $\frac{1}{2}$, which is not zero, the functions $p(x)$ and $q(x)$ are not orthogonal.