Lecture 17: Linear Transformations

MAT2040 Linear Algebra

Recall that when we defined the matrix product AB, I tried to convince you the definition was logical by thinking about the composition of two *mappings*.

Recall that when we defined the matrix product AB, I tried to convince you the definition was logical by thinking about the composition of two *mappings*.

We will now commence the study of transformations like $x \mapsto A\mathbf{x}$. In particular we will generalize the notion of transformations, and think about transformations from *one vector space to another vector space*.

Definition 17.1

$L(\cdot)$ is a **mapping** (or **transformation**) from vector space V to vector space W, if L assigns a value (vector) in W for every vector $\mathbf{v} \in V$.

Definition 17.1

$L(\cdot)$ is a **mapping** (or **transformation**) from vector space V to vector space W, if L assigns a value (vector) in W for every vector $\mathbf{v} \in V$.

We write $\mathbf{v} \mapsto L(\mathbf{v})$.

Definition 17.1

$L(\cdot)$ is a **mapping** (or **transformation**) from vector space V to vector space W, if L assigns a value (vector) in W for every vector $\mathbf{v} \in V$.

We write $\mathbf{v} \mapsto L(\mathbf{v})$.

Example: mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$: $\mathbf{x} \mapsto A\mathbf{x}$ (where A is an $m \times n$ matrix).

For the mapping $\mathbf{x} \mapsto A\mathbf{x}$ (from $\mathbb{R}^n$ to $\mathbb{R}^m$) we defined the following concepts:

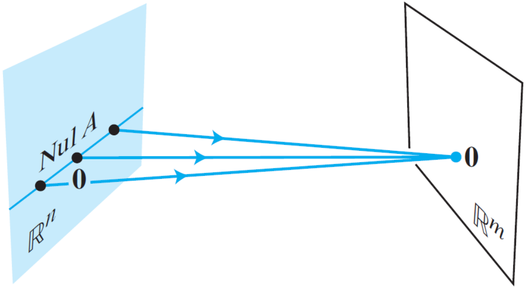
- $\boxed{}$ set of all possible values of $\mathbb{R}^m$ this mapping can achieve,
- $\boxed{}$ set of all possible value of $\mathbb{R}^n$ that are mapped to $\mathbf{0}$.

For the mapping $\mathbf{x} \mapsto A\mathbf{x}$ (from $\mathbb{R}^n$ to $\mathbb{R}^m$) we defined the following concepts:

- $\boxed{\text{Col } A =}$ set of all possible values of $\mathbb{R}^m$ this mapping can achieve,
- $\boxed{\phantom{\text{Nul } A =}}$ set of all possible value of $\mathbb{R}^n$ that are mapped to $\mathbf{0}$.

For the mapping $\mathbf{x} \mapsto A\mathbf{x}$ (from $\mathbb{R}^n$ to $\mathbb{R}^m$) we defined the following concepts:

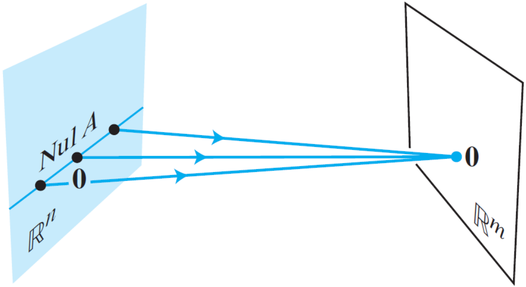
- $\boxed{\text{Col } A =}$ set of all possible values of $\mathbb{R}^m$ this mapping can achieve,
- $\boxed{\text{Null } A =}$ set of all possible value of $\mathbb{R}^n$ that are mapped to $\mathbf{0}$.



(picture from David Lay, Linear Algebra.)

For the mapping $\mathbf{x} \mapsto A\mathbf{x}$ (from $\mathbb{R}^n$ to $\mathbb{R}^m$) we defined the following concepts:

- $\boxed{\text{Col } A =}$ set of all possible values of $\mathbb{R}^m$ this mapping can achieve,
- $\boxed{\text{Null } A =}$ set of all possible value of $\mathbb{R}^n$ that are mapped to $\mathbf{0}$.



(picture from David Lay, Linear Algebra.)

We want to have words to refer to these concepts for general mappings from V to W as well!

Definition 17.2

For the mapping $\mathbf{v} \mapsto L(v)$ (from V to W) we define:

- $\boxed{}$ set of all possible values of W this mapping can achieve,
- $\boxed{}$ set of all possible value of V that are mapped to $\mathbf{0} \in W$.

Definition 17.2

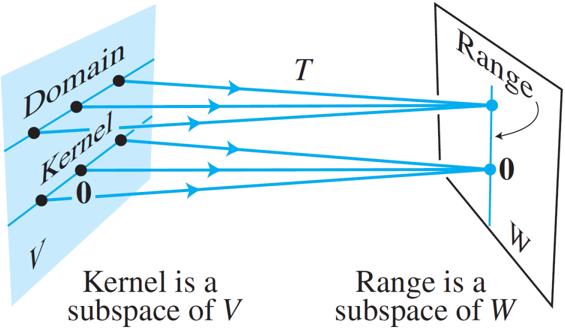
For the mapping $\mathbf{v} \mapsto L(v)$ (from V to W) we define:

- $\boxed{\textbf{the range of } L =}$ set of all possible values of W this mapping can achieve,
- $\boxed{\phantom{\text{the kernel of } L =}}$ set of all possible value of V that are mapped to $\mathbf{0} \in W$.

Definition 17.2

For the mapping $\mathbf{v} \mapsto L(v)$ (from V to W) we define:

- $\boxed{\textbf{the range of } L =}$ set of all possible values of W this mapping can achieve,
- $\boxed{\textbf{the kernel of } L =}$ set of all possible value of V that are mapped to $\mathbf{0} \in W$.



(picture from David Lay, Linear Algebra.)

Example 17.3

Consider the mapping $f(x)$ from $\mathbb{R}$ to $\mathbb{R}$ given by $x \mapsto \max\{x, 0\}$ (the absolute value of x).

What is the range of $f(x)$? What is the kernel of $f(x)$?

Example 17.3

Consider the mapping $f(x)$ from $\mathbb{R}$ to $\mathbb{R}$ given by $x \mapsto \max\{x, 0\}$ (the absolute value of x).

What is the range of $f(x)$? What is the kernel of $f(x)$?

Consider $L(\mathbf{x}) = \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix} \mathbf{x}$.

What is the range of $L(\mathbf{x})$? What is the kernel of $L(\mathbf{x})$?

Not every mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$ can be expressed as $x \mapsto A\mathbf{x}$ for some matrix A.

Not every mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$ can be expressed as $x \mapsto A\mathbf{x}$ for some matrix A.

For instance the mapping $x \mapsto x^2$ (from $\mathbb{R}$ to $\mathbb{R}$) cannot be expressed as a matrix multiplication.

The following definition summarizes which transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$ **can** be expressed as $x \mapsto A\mathbf{x}$ for some matrix A.

The following definition summarizes which transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$ **can** be expressed as $x \mapsto A\mathbf{x}$ for some matrix A.

Definition 17.4

A mapping $L(\cdot)$ from $\mathbb{R}^n$ to $\mathbb{R}^m$ is called **a linear transformation**, if

$$L(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha L(\mathbf{x}) + \beta L(\mathbf{y})$$

for every $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\alpha, \beta \in \mathbb{R}$.

The following definition summarizes which transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$ **can** be expressed as $x \mapsto A\mathbf{x}$ for some matrix A.

Definition 17.4

A mapping $L(\cdot)$ from $\mathbb{R}^n$ to $\mathbb{R}^m$ is called **a linear transformation**, if

$$L(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha L(\mathbf{x}) + \beta L(\mathbf{y})$$

for every $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\alpha, \beta \in \mathbb{R}$.

In other words: every linear combination of vectors in $\mathbb{R}^n$ is mapped to the linear combination of the image of these vectors with the same weights.

Lemma 17.5

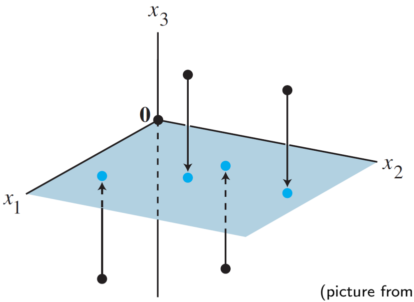
$L(\mathbf{x})$ from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a linear transformation, if and only if $L(\mathbf{x}) = A\mathbf{x}$ for some $m \times n$ matrix A.

Example 17.6 (Linear Transformation)

$$\mathbf{x} \mapsto A\mathbf{x} \text{ where } A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Example 17.6 (Linear Transformation)

$\mathbf{x} \mapsto A\mathbf{x}$ where $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.



(picture from David Lay, Linear Algebra.)

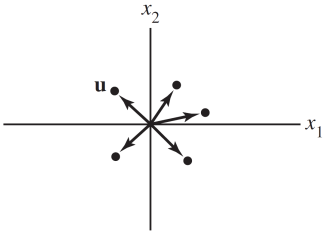
Projection Transformation

Example 17.7 (Linear Transformation)

$T(\mathbf{x}) = t\mathbf{x}$, where $t > 1$.

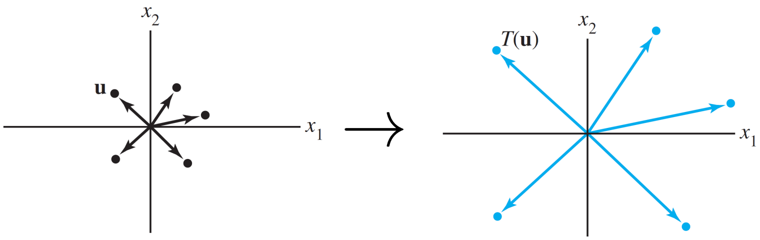
Example 17.7 (Linear Transformation)

$T(\mathbf{x}) = t\mathbf{x}$, where $t > 1$.



Example 17.7 (Linear Transformation)

$T(\mathbf{x}) = t\mathbf{x}$, where $t > 1$.



(picture from David Lay, Linear Algebra.)

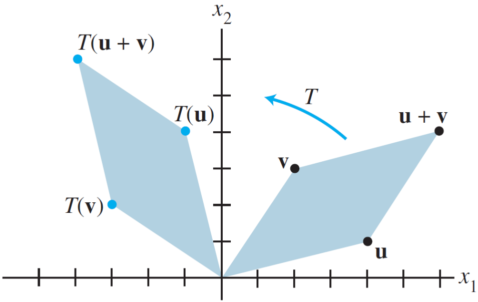
Dilation Transformation

Example 17.8 (Linear Transformation)

$T(\mathbf{x}) = A\mathbf{x}$ where $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

Example 17.8 (Linear Transformation)

$T(\mathbf{x}) = A\mathbf{x}$ where $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.



(picture from David Lay, Linear Algebra.)

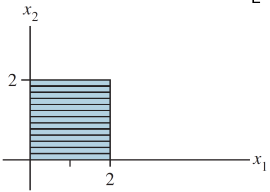
Rotation Transformation

Example 17.9 (Linear Transformation)

$T(\mathbf{x}) = A\mathbf{x}$ where $A = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}$.

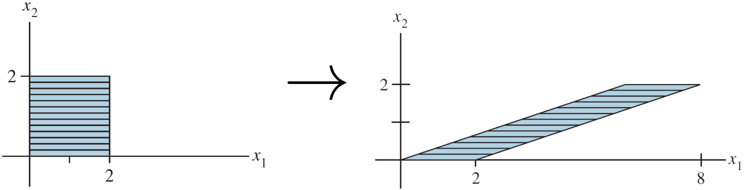
Example 17.9 (Linear Transformation)

$T(\mathbf{x}) = A\mathbf{x}$ where $A = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}$.



Example 17.9 (Linear Transformation)

$T(\mathbf{x}) = A\mathbf{x}$ where $A = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}$.



(picture from David Lay, Linear Algebra.)

Shear Transformation

Key takeaway: If we have

- a linear transformation L from $\mathbb{R}^n$ to $\mathbb{R}^m$,
- a basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ for $\mathbb{R}^n$, and
- we know $L(\mathbf{b}_1), L(\mathbf{b}_2), \ldots, L(\mathbf{b}_n)$,

then we know $L(\mathbf{x})$ for any $\mathbf{x} \in \mathbb{R}^n$:

Key takeaway: If we have

- a linear transformation L from $\mathbb{R}^n$ to $\mathbb{R}^m$,
- a basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ for $\mathbb{R}^n$, and
- we know $L(\mathbf{b}_1), L(\mathbf{b}_2), \ldots, L(\mathbf{b}_n)$,

then we know $L(\mathbf{x})$ for any $\mathbf{x} \in \mathbb{R}^n$: we know that we can write

$$\mathbf{x} = t_1\mathbf{b}_1 + t_2\mathbf{b}_2 + \cdots + t_n\mathbf{b}_n$$

for some $t_1, t_2, \ldots, t_n \in \mathbb{R}$ (because $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n$ is a basis for $\mathbb{R}^n$),

Key takeaway: If we have

- a linear transformation L from $\mathbb{R}^n$ to $\mathbb{R}^m$,
- a basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ for $\mathbb{R}^n$, and
- we know $L(\mathbf{b}_1), L(\mathbf{b}_2), \ldots, L(\mathbf{b}_n)$,

then we know $L(\mathbf{x})$ for any $\mathbf{x} \in \mathbb{R}^n$: we know that we can write

$$\mathbf{x} = t_1\mathbf{b}_1 + t_2\mathbf{b}_2 + \cdots + t_n\mathbf{b}_n$$

for some $t_1, t_2, \ldots, t_n \in \mathbb{R}$ (because $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n$ is a basis for $\mathbb{R}^n$), and so by linearity of L we get:

Key takeaway: If we have

- a linear transformation L from $\mathbb{R}^n$ to $\mathbb{R}^m$,
- a basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ for $\mathbb{R}^n$, and
- we know $L(\mathbf{b}_1), L(\mathbf{b}_2), \ldots, L(\mathbf{b}_n)$,

then we know $L(\mathbf{x})$ for any $\mathbf{x} \in \mathbb{R}^n$: we know that we can write

$$\mathbf{x} = t_1\mathbf{b}_1 + t_2\mathbf{b}_2 + \cdots + t_n\mathbf{b}_n$$

for some $t_1, t_2, \ldots, t_n \in \mathbb{R}$ (because $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n$ is a basis for $\mathbb{R}^n$), and so by linearity of L we get:

$$L(\mathbf{x}) = t_1 L(\mathbf{b}_1) + t_2 L(\mathbf{b}_2) + \cdots + t_n L(\mathbf{b}_n).$$

Key takeaway: If we have

- a linear transformation L from $\mathbb{R}^n$ to $\mathbb{R}^m$,
- a basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ for $\mathbb{R}^n$, and
- we know $L(\mathbf{b}_1), L(\mathbf{b}_2), \ldots, L(\mathbf{b}_n)$,

then we know $L(\mathbf{x})$ for any $\mathbf{x} \in \mathbb{R}^n$: we know that we can write

$$\mathbf{x} = t_1\mathbf{b}_1 + t_2\mathbf{b}_2 + \cdots + t_n\mathbf{b}_n$$

for some $t_1, t_2, \ldots, t_n \in \mathbb{R}$ (because $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n$ is a basis for $\mathbb{R}^n$), and so by linearity of L we get:

$$L(\mathbf{x}) = t_1L(\mathbf{b}_1) + t_2L(\mathbf{b}_2) + \cdots + t_nL(\mathbf{b}_n).$$

Note that $[\mathbf{x}]_{\mathcal{B}} = (t_1, t_2, \ldots, t_n)^T$!

Key takeaway: If we have

- a linear transformation L from $\cancel{\mathbb{R}^n}$ a vector space V to $\cancel{\mathbb{R}^m}$ a vector space W,
- a basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ for $\cancel{\mathbb{R}^n}$ V, and
- we know $L(\mathbf{b}_1), L(\mathbf{b}_2), \ldots, L(\mathbf{b}_n)$,

then we know $L(\mathbf{x})$ for any $\mathbf{x} \in \cancel{\mathbb{R}^n}$ V: we know that we can write

$$\mathbf{x} = t_1\mathbf{b}_1 + t_2\mathbf{b}_2 + \cdots + t_n\mathbf{b}_n$$

for some $t_1, t_2, \ldots, t_n \in \mathbb{R}$ (because $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n$ is a basis for $\cancel{\mathbb{R}^n}$ V), and so by linearity of L we get:

$$L(\mathbf{x}) = t_1 L(\mathbf{b}_1) + t_2 L(\mathbf{b}_2) + \cdots + t_n L(\mathbf{b}_n).$$

Note that $[\mathbf{x}]_{\mathcal{B}} = (t_1, t_2, \ldots, t_n)^T$!

Definition 17.8 (for general vector spaces)

A mapping $L(\cdot)$ from V to W is called **a linear transformation**, if

$$L(\alpha \mathbf{x} + \beta \mathbf{y}) = \alpha L(\mathbf{x}) + \beta L(\mathbf{y})$$

for every $\mathbf{x}, \mathbf{y} \in V$ and $\alpha, \beta \in \mathbb{R}$.

Example 17.10

Consider the vector space P_k (polynomials of degree at most k).

Is *differentiation* a linear transformation?

Example 17.11

Let V be a vector space with a basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$.

Is $L(\mathbf{x}) = [\mathbf{x}]_{\mathcal{B}}$ a linear transformation?

Example 17.11

Let V be a vector space with a basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$.

Is $L(\mathbf{x}) = [\mathbf{x}]_{\mathcal{B}}$ a linear transformation?

Recall:

Lemma 12.6

Let V be a vector space with a basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$, and let $\mathbf{x}, \mathbf{y} \in V$. For any $\alpha, \beta \in \mathbb{R}$

$$[\alpha\mathbf{x} + \beta\mathbf{y}]_{\mathcal{B}} = \alpha[\mathbf{x}]_{\mathcal{B}} + \beta[\mathbf{y}]_{\mathcal{B}}.$$

Theorem 17.12

Let V be a vector space with basis $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ and W be a vector space with basis $\mathcal{C} = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$.
If $T(\mathbf{x})$ from V to W is a linear transformation then there exists an $m \times n$ matrix A so that

$$[T(\mathbf{x})]_{\mathcal{C}} = A[\mathbf{x}]_{\mathcal{B}}.$$

Theorem 17.12

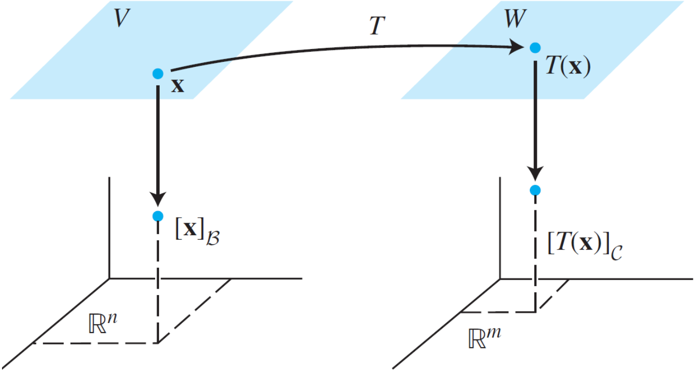
Let V be a vector space with basis $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ and W be a vector space with basis $\mathcal{C} = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$.
If $T(\mathbf{x})$ from V to W is a linear transformation then there exists an $m \times n$ matrix A so that

$$[T(\mathbf{x})]_{\mathcal{C}} = A[\mathbf{x}]_{\mathcal{B}}.$$

In fact,

$$A = \left[[T(\mathbf{v}_1)]_{\mathcal{C}}, \quad [T(\mathbf{v}_2)]_{\mathcal{C}}, \quad \ldots, [T(\mathbf{v}_n)]_{\mathcal{C}}\right].$$

A is called the **matrix for T relative to the bases $\mathcal{B}$ and $\mathcal{C}$**.



(picture from David Lay, Linear Algebra.)

The last theorem is saying that every linear transformation T from a vector space V to a vector space W can be represented (and thus studied) by a matrix!

The last theorem is saying that every linear transformation T from a vector space V to a vector space W can be represented (and thus studied) by a matrix!

The trick is the following:

- determine a basis $\mathcal{B}$ for V, and $\mathcal{C}$ for W
- for every $\mathbf{v}_i \in \mathcal{B}$, determine $T(\mathbf{v}_i)$
- the **coordinates** of $T(\mathbf{v}_i)$ with respect to basis $\mathcal{C}$ are the i-th column vector of A.

The last theorem is saying that every linear transformation T from a vector space V to a vector space W can be represented (and thus studied) by a matrix!

The trick is the following:

- determine a basis $\mathcal{B}$ for V, and $\mathcal{C}$ for W
- for every $\mathbf{v}_i \in \mathcal{B}$, determine $T(\mathbf{v}_i)$
- the **coordinates** of $T(\mathbf{v}_i)$ with respect to basis $\mathcal{C}$ are the i-th column vector of A.

Linearity makes sure (the $\mathcal{B}$-coordinates) of all other vectors in V are mapped to the correct $\mathcal{C}$-coordinates in W.

Example 17.13

Integration is a linear transformation from P_k to P_{k+1}.

Example 17.13

Integration is a linear transformation from P_k to P_{k+1}.

Choose a basis for P_k and P_{k+1}, and find the matrix for integration relative to your chosen bases.