0.1 Linear Transformations

A *function* is a rule that assigns a value from a set $B$ for each element in a set $A$.

Notation: $f : A \mapsto B$

If the value $b \in B$ is assigned to value $a \in A$, then write $f(a) = b$, $b$ is called the *image* of $a$ under $f$.

$A$ is called the *domain* of $f$ and $B$ is called the *codomain*.

The subset of $B$ consisting of all possible values of $f$ as $a$ varies in the domain is called the *range* of $f$.

**Definition 1**

Two functions $f_1$ and $f_2$ are called equal, if their domains are equal and $f_1(a) = f_2(a)$ for all $a$ in the domain.

**Example 1**

<table>
<thead>
<tr>
<th>function</th>
<th>example</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f(x)$</td>
<td>$f(x) = x - 2$</td>
<td>Function from $\mathbb{R}$ to $\mathbb{R}$</td>
</tr>
<tr>
<td>$f(x,y)$</td>
<td>$f(x,y) = x + y$</td>
<td>Function from $\mathbb{R}^2$ to $\mathbb{R}$</td>
</tr>
<tr>
<td>$f(x,y,z)$</td>
<td>$f(x,y,z) = x + y + z$</td>
<td>Function from $\mathbb{R}^3$ to $\mathbb{R}$</td>
</tr>
<tr>
<td>$f(x,y,z)$</td>
<td>$f(x,y,z) = (x + y, z)$</td>
<td>Function from $\mathbb{R}^3$ to $\mathbb{R}^2$</td>
</tr>
</tbody>
</table>

**Functions from** $\mathbb{R}^n$ **to** $\mathbb{R}^m$ If $f : \mathbb{R}^n \mapsto \mathbb{R}^m$, then $f$ is called a *map* or a *transformation*.

If $m = n$, then $f$ is called an operator on $\mathbb{R}^n$.

Let $f_1, f_2, \ldots, f_m$ functions from $\mathbb{R}^n$ to $\mathbb{R}$, assume

$$
\begin{align*}
    f_1(x_1, x_2, \ldots, x_n) &= w_1 \\
    f_2(x_1, x_2, \ldots, x_n) &= w_2 \\
    \vdots \\
    f_m(x_1, x_2, \ldots, x_n) &= w_m
\end{align*}
$$

then the point $(w_1, w_2, \ldots, w_m) \in \mathbb{R}^m$ is assigned to $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ and thus those functions define a transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$.

Denote the transformation $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ and

$$
T(x_1, x_2, \ldots, x_n) = (w_1, w_2, \ldots, w_m)
$$

**Example 2**

$f_1(x_1, x_2) = x_1 + x_2, f_2(x_1, x_2) = x_1 x_2$ define an operator $T : \mathbb{R}^2 \mapsto \mathbb{R}^2$.

$$
T(x_1, x_2) = (x_1 + x_2, x_1 x_2)
$$

**Linear Transformations** In the special case where the functions $f_1, f_2, \ldots, f_m$ are linear, the transformation $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ is called a linear transformation.
A linear transformation is defined by equations

\[
\begin{align*}
 a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= w_1 \\
 a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= w_2 \\
 \vdots & \vdots \\
 a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= w_m
\end{align*}
\]

or in matrix notation

\[
\begin{pmatrix}
 a_{11} & a_{12} & \ldots & a_{1n} \\
 a_{21} & a_{22} & \ldots & a_{2n} \\
 \vdots & \vdots & \ddots & \vdots \\
 a_{m1} & a_{m2} & \ldots & a_{mn}
\end{pmatrix}
\begin{pmatrix}
 x_1 \\
 x_2 \\
 \vdots \\
 x_n
\end{pmatrix}
= 
\begin{pmatrix}
 w_1 \\
 w_2 \\
 \vdots \\
 w_m
\end{pmatrix}
\]

or

\[Ax = w\]

The matrix \(A\) is called the standard matrix for the linear transformation \(T\), and \(T\) is called multiplication by \(A\).

**Remark:**
Through this discussion we showed that a linear transformation from \(\mathbb{R}^n\) to \(\mathbb{R}^m\) correspond to matrices of size \(m \times n\).

One can say that to each matrix \(A\) there corresponds a linear transformation \(T : \mathbb{R}^n \mapsto \mathbb{R}^m\), and to each linear \(T : \mathbb{R}^n \mapsto \mathbb{R}^m\) transformation there corresponds an \(m \times n\) matrix \(A\).

**Example 3**
Let \(T : \mathbb{R}^3 \mapsto \mathbb{R}^2\) defined by

\[
\begin{align*}
 2x_1 + 3x_2 + (-1)x_3 &= w_1 \\
  x_1 + x_2 + (-1)x_3 &= w_2
\end{align*}
\]

can be expressed in matrix form as

\[
\begin{pmatrix}
 2 & 3 & -1 \\
 1 & 1 & -1
\end{pmatrix}
\begin{pmatrix}
  x_1 \\
  x_2 \\
  x_3
\end{pmatrix}
= 
\begin{pmatrix}
  w_1 \\
  w_2
\end{pmatrix}
\]

The standard matrix for \(T\) is

\[
\begin{pmatrix}
  2 & 3 & -1 \\
  1 & 1 & -1
\end{pmatrix}
\]

The image of a point \((x_1, x_2, x_3)\) can be found by using the defining equations or by matrix multiplication.

\[
T(1, 2, 0) = \begin{pmatrix}
  2 & 3 & -1 \\
  1 & 1 & -1
\end{pmatrix}
\begin{pmatrix}
  1 \\
  2 \\
  0
\end{pmatrix}
= \begin{pmatrix}
  8 \\
  3
\end{pmatrix}
\]
Notation:
If \( T : \mathbb{R}^n \to \mathbb{R}^m \) is a multiplication by \( A \), and if it important to emphasize the standard matrix then we shall denote the transformation by \( T_A : \mathbb{R}^n \to \mathbb{R}^m \). Thus
\[
T_A(x) = Ax
\]
Since linear transformations can be identified with their standard matrices we will use \([T]\) as symbol for the standard matrix for \( T : \mathbb{R}^n \to \mathbb{R}^m \).
\[
T(x) = [T]x \text{ or } [T_A] = A
\]

Geometry of linear Transformations
A linear transformation \( T : \mathbb{R}^n \to \mathbb{R}^m \) transforms points in \( \mathbb{R}^n \) into new points in \( \mathbb{R}^m \)

Example 4
Zero Transformation The zero transformation from \( T_0 : \mathbb{R}^n \to \mathbb{R}^m \) has standard matrix 0, so that
\[
T_0(x) = 0
\]
for all \( x \in \mathbb{R}^n \)

Example 5
Identity Transformation The identity transformation \( T_I : \mathbb{R}^n \to \mathbb{R}^m \) has standard matrix \( I_n \) (\( n \times n \) identity matrix), so that
\[
T_I(x) = I_n x = x
\]
for all \( x \in \mathbb{R}^n \).

Among the more important transformations are those that cause reflections, projections, and rotations

Example 6
Reflections
Consider \( T : \mathbb{R}^2 \to \mathbb{R}^2 \) with standard matrix
\[
\begin{bmatrix}
-1 & 0 \\
0 & 1
\end{bmatrix}
\]
then
\[
T(x) = \begin{bmatrix}
-1 & 0 \\
0 & 1
\end{bmatrix} x = \begin{bmatrix}
-x_1 \\
x_2
\end{bmatrix}
\]
\( T \) reflects points \((x_1, x_2)\) about the \( y \)-axis.
What might be the standard matrix of the linear transformation reflecting point about the \( x \)-axis?
\[
\mathbb{R}^2 \to \mathbb{R}^2
\]
Reflection about the $y$-axis  \( T(x, y) = (-x, y) \) \[
\begin{bmatrix}
-1 & 0 \\
0 & 1 \\
\end{bmatrix}
\]

Reflection about the $x$-axis  \( T(x, y) = (x, -y) \) \[
\begin{bmatrix}
1 & 0 \\
0 & -1 \\
\end{bmatrix}
\]

Reflection about the line $y = x$  \( T(x, y) = (y, x) \) \[
\begin{bmatrix}
0 & 1 \\
1 & 0 \\
\end{bmatrix}
\]

Reflection about the $xy$-plane  \( T(x, y, z) = (x, y, -z) \) \[
\begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & -1 \\
\end{bmatrix}
\]

Reflection about the $xz$-plane  \( T(x, y, z) = (x, -y, z) \) \[
\begin{bmatrix}
1 & 0 & 0 \\
0 & -1 & 0 \\
0 & 0 & 1 \\
\end{bmatrix}
\]

Reflection about the $yz$-planes  \( T(x, y, z) = (-x, y, z) \) \[
\begin{bmatrix}
-1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1 \\
\end{bmatrix}
\]

Example 7  
**Projections** Consider \( T : \mathbb{R}^2 \mapsto \mathbb{R}^2 \) with standard matrix \[
\begin{bmatrix}
1 & 0 \\
0 & 0 \\
\end{bmatrix}
\]
then \( T(x) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} x = \begin{bmatrix} x_1 \\ 0 \end{bmatrix} \)
It gives the orthogonal projection of point \((x, y)\) onto the $x$-axis.

Consider \( T : \mathbb{R}^3 \mapsto \mathbb{R}^3 \) with standard matrix \[
\begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 0 \\
\end{bmatrix}
\]
then \( T(x) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} x = \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix} \)
It gives the orthogonal projection of a point \((x, y, z)\) onto the $xy$-plane.
Example 8
Rotation:
An operator that rotates a vector in \( \mathbb{R}^2 \) through a given angle \( \theta \) is called a rotation operator in \( \mathbb{R}^2 \).

\[
T_R(x) = (x \cos \theta - y \sin \theta, x \sin \theta + y \cos \theta)
\]
The standard matrix of a rotation operator in \( \mathbb{R}^2 \) for angle \( \theta \) is therefore

\[
[T_R] = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}
\]

Proof:

Let \((w_1, w_2) = T_R(x)\), then (check the diagram)

\[
w_1 = r \cos(\theta + \varphi), \quad w_2 = r \sin(\theta + \varphi)
\]
Using trigonometric identities
\[ w_1 = r \cos(\theta) \cos(\varphi) - r \sin(\theta) \sin(\varphi) \]
\[ w_2 = r \sin(\theta) \cos(\varphi) + r \cos(\theta) \sin(\varphi) \]

Also (check diagram)
\[ x = r \cos \varphi, \quad y = r \sin \varphi \]

Substituting the later into the equations above gives
\[ w_1 = x \cos(\theta) - y \sin(\theta), \quad w_2 = x \sin(\theta) + y \cos(\theta) \]

Therefore
\[ T_R(x, y) = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \]

**Example 9**

The standard matrix of the rotation by \( \pi/2 \) is
\[ [T_R] = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \]

Therefore
\[ T_R(1, 2) = \begin{bmatrix} -2 \\ 1 \end{bmatrix} \]

The standard matrix of the rotation by \( \pi/4 \) is
\[ [T_R] = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \]

Therefore
\[ T_R(1, 2) = \begin{bmatrix} -1/\sqrt{2} \\ 3/\sqrt{2} \end{bmatrix} \]

**Rotation in \( \mathbb{R}^3 \)**

<table>
<thead>
<tr>
<th>Operator</th>
<th>Equation</th>
<th>Standard matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>Counterclockwise rotation about the positive ( x )-axis through an angle ( \theta )</td>
<td>( T(x, y, z) = \begin{bmatrix} x \ y \cos \theta - z \sin \theta \ y \sin \theta + z \cos \theta \end{bmatrix} )</td>
<td>( \begin{bmatrix} 1 &amp; 0 &amp; 0 \ 0 &amp; \cos \theta &amp; -\sin \theta \ 0 &amp; \sin \theta &amp; \cos \theta \end{bmatrix} )</td>
</tr>
<tr>
<td>Counterclockwise rotation about the positive ( y )-axis through an angle ( \theta )</td>
<td>( T(x, y, z) = \begin{bmatrix} x \cos \theta + z \sin \theta \ y \ -x \sin \theta + z \cos \theta \end{bmatrix} )</td>
<td>( \begin{bmatrix} \cos \theta &amp; 0 &amp; \sin \theta \ 0 &amp; 1 &amp; 0 \ -\sin \theta &amp; 0 &amp; \cos \theta \end{bmatrix} )</td>
</tr>
<tr>
<td>Counterclockwise rotation about the positive ( z )-axis through an angle ( \theta )</td>
<td>( T(x, y, z) = \begin{bmatrix} x \cos \theta - y \sin \theta \ x \sin \theta + y \cos \theta \ z \end{bmatrix} )</td>
<td>( \begin{bmatrix} \cos \theta &amp; -\sin \theta &amp; 0 \ \sin \theta &amp; \cos \theta &amp; 0 \ 0 &amp; 0 &amp; 1 \end{bmatrix} )</td>
</tr>
</tbody>
</table>
Dilation and Contraction
This is the operator stretching or shrinking a vector by a factor $k$, but keeping the direction un-
changed. We call the operator a dilation if the transformed vector is at least as long as the original
vector, and a contraction if the transformed vector is at most as long as the original vector.

<table>
<thead>
<tr>
<th>Operator</th>
<th>Equation</th>
<th>Standard matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contraction with factor $k$ on $\mathbb{R}^2$, $0 \leq k \leq 1$</td>
<td>$T(x, y) = \begin{bmatrix} kx \ ky \end{bmatrix}$</td>
<td>$\begin{bmatrix} k &amp; 0 \ 0 &amp; k \end{bmatrix}$</td>
</tr>
<tr>
<td>Dilation with factor $k$ on $\mathbb{R}^2$, $k \geq 1$</td>
<td>$T(x, y) = \begin{bmatrix} kx \ ky \end{bmatrix}$</td>
<td>$\begin{bmatrix} k &amp; 0 \ 0 &amp; k \end{bmatrix}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Operator</th>
<th>Equation</th>
<th>Standard matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contraction with factor $k$ on $\mathbb{R}^3$, $0 \leq k \leq 1$</td>
<td>$T(x, y, z) = \begin{bmatrix} kx \ ky \ kz \end{bmatrix}$</td>
<td>$\begin{bmatrix} k &amp; 0 &amp; 0 \ 0 &amp; k &amp; 0 \ 0 &amp; 0 &amp; k \end{bmatrix}$</td>
</tr>
<tr>
<td>Dilation with factor $k$ on $\mathbb{R}^3$, $k \geq 1$</td>
<td>$T(x, y, z) = \begin{bmatrix} kx \ ky \ kz \end{bmatrix}$</td>
<td>$\begin{bmatrix} k &amp; 0 &amp; 0 \ 0 &amp; k &amp; 0 \ 0 &amp; 0 &amp; k \end{bmatrix}$</td>
</tr>
</tbody>
</table>

Composition of Linear Transformations
Let $T_A : \mathbb{R}^n \mapsto \mathbb{R}^k$ and $T_B : \mathbb{R}^k \mapsto \mathbb{R}^m$ be linear transformations, then for each $x \in \mathbb{R}^n$ one can first compute $T_A(x)$, which is a vector in $\mathbb{R}^k$ and then one can compute $T_B(T_A(x))$, which is a vector in $\mathbb{R}^m$.

Thus the application of first $T_A$ and then of $T_B$ is a transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$. It is called the composition of $T_B$ with $T_A$, and is denoted as $T_B \circ T_A$ (read $T_B$ circle $T_A$).

$$(T_B \circ T_A)(x) = T_B(T_A(x)) = T_B(Ax) = B(Ax) = BAx$$

Therefore the standard matrix of the composition of $T_B$ with $T_A$ is $BA$.

$$T_B \circ T_A = T_{BA}$$

**Remark:**
This equation points out an important interpretation of the matrix product. Composition of two linear transformations is equivalent to the multiplication of two matrices.

**Example 10**
**In general:** Composition is not commutative.
$T_1$: reflection about $y = x$, and $T_2$ orthogonal projection onto $y$
One can easily generalize the concept to the composition of more than two transformations.
0.2 Properties of linear Transformations

One-to-One Linear Transformations
Transformations that transform different vectors into different images, that is
If \( x \neq y \) therefore \( T(x) \neq T(y) \),
are of special interest.
One such example is the rotation by an angle \( \theta \) in \( \mathbb{R}^2 \).
But the orthogonal projection onto the \( xy - plane \) in \( \mathbb{R}^3 \) does not have this property.
\( T(x_1, x_2, x_3) = (x_1, x_2, 0) \), so \( T(2, 1, 1) = T(2, 1, 45) \).

Definition 2
A linear transformation \( T : \mathbb{R}^n \mapsto \mathbb{R}^m \) is said to be one-to-one, if it is true that
\[
x \neq y \Rightarrow T(x) \neq T(y)
\]
distinct vectors in \( \mathbb{R}^n \) are mapped into distinct vectors in \( \mathbb{R}^m \).

Conclusion:
If \( T \) is one-to-one and \( w \) is a vector in the range of \( T \), then there is exactly one vector in \( \mathbb{R}^n \) with
\( T(x) = w \).

Consider transformations from \( \mathbb{R}^n \) to \( \mathbb{R}^n \), then the standard matrices are square matrices of size \( n \times n \).

Theorem 1
If \( A \) is a \( n \times n \) matrix and \( T_A : \mathbb{R}^n \mapsto \mathbb{R}^n \) is the multiplication by \( A \), then the following statements
are equivalent
(a) \( A \) is invertible
(b) The range of \( T_A \) is \( \mathbb{R}^n \)
(c) \( T_A \) is one-to-one.

Proof:
(a)\( \Rightarrow \) (b) 
(b)\( \Rightarrow \) (c) 
(c)\( \Rightarrow \) (a) 

Application:
The rotation by \( \theta \) in \( \mathbb{R}^2 \) is one-to-one.
The standard matrix of this operator is
\[
A = \begin{bmatrix}
\cos \theta & -\sin \theta \\
\sin \theta & \cos \theta
\end{bmatrix}
\]
then \( \det(A) = \cos^2 \theta + \sin^2 \theta = 1 \neq 0 \), therefore \( A \) is invertible and therefore the rotation operator is one-to-one.

Show yourself using this criteria that the orthogonal projection in \( \mathbb{R}^3 \) is NOT one-to-one.

**Inverse of a one-to-one Operator**

**Definition 3**
If \( T_A : \mathbb{R}^n \mapsto \mathbb{R}^n \) is a one-to-one operator, then \( T^{-1} = T_{A^{-1}} \) is called the inverse operator of \( T_A \).

**Example 11**
Let \( T : \mathbb{R}^2 \mapsto \mathbb{R}^2 \) with

\[
[T] = \begin{bmatrix} 1 & -1 \\ -2 & 0 \end{bmatrix}
\]

Then

\[
T(x_1, x_2) = (x_1 - x_2, -2x_1) \text{ for } (x_1, x_2) \in \mathbb{R}^2
\]

and since

\[
[T]^{-1} = -\frac{1}{2} \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}
\]

and

\[
T^{-1}(x_1, x_2) = (-x_2/2, -(2x_1 + x_2)/2).
\]

**Theorem 2**
Let \( T : \mathbb{R}^n \mapsto \mathbb{R}^n \) be a one-to-one operator, then

(a) \[
T(x) = w \iff T^{-1}(w) = x
\]

(b) For \( x \in \mathbb{R}^n \) it is \( (T \circ T^{-1})(x) = x \), and \( (T^{-1} \circ T)(x) = x \)

**Proof:**

(a) If \( T : \mathbb{R}^n \mapsto \mathbb{R}^n \) is a one-to-one operator and \( T(x) = w \), then the standard matrix \([T]\) is invertible and

\[
\begin{align*}
T(x) &= w \\
\iff [T]x &= w \\
\iff [T]^{-1}[T]x &= [T]^{-1}w \\
\iff I_n x &= [T]^{-1}w \\
\iff x &= [T^{-1}]w \\
\iff x &= T^{-1}(w)
\end{align*}
\]
(b) If $T : \mathbb{R}^n \mapsto \mathbb{R}^n$ is a one-to-one operator and $T(x) = w$, then the standard matrix $A$ is invertible and

$$T \circ T^{-1} = T_A \circ T_{A^{-1}} = T_{AA^{-1}} = T_I$$

therefore the claim holds.

**Theorem 3**

A transformation $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ is linear if and only if the following relationship holds for all vectors $u, v \in \mathbb{R}^n$ and $c \in \mathbb{R}$.

1. $T(u + v) = T(u) + T(v)$
2. $T(cu) = cT(u)$

**Example 12**

Show that $T(x_1, x_2) = (x_1 - x_2, -2x_1)$ for $(x_1, x_2) \in \mathbb{R}^2$ is a linear transformation.

(a) Let $u, v \in \mathbb{R}^2$, then

$$T(u + v) = (u_1 + v_1 - (u_2 + v_2), -2(u_1 + v_1)) = (u_1 - u_2, -2u_1) + (v_1 - v_2, -2v_1) = T(u) + T(v)$$

(b)

$$T(cu) = (cu_1 - (cu_2)), -2(cu_1)) = c(u_1 - u_2, -2u_1) = cT(u)$$

Both properties hold, therefore $T$ is a linear transformation.

**Theorem 4**

If $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ is a linear transformation and $e_1, e_2, \ldots, e_n$ are the standard basis vectors for $\mathbb{R}^n$, then the standard matrix for $T$ is

$$[T] = \begin{bmatrix} T(e_1) & T(e_2) & \cdots & T(e_n) \end{bmatrix}$$
Example 13
Let $T$ be the orthogonal projection onto the $yz$-plane in $\mathbb{R}^3$. Then

$$T(e_1) = T\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 0 \end{bmatrix},
T(e_2) = T\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 1 \end{bmatrix},
T(e_3) = T\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

and therefore

$$[T] = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$