4. Lecture 4: 2.1 Linear Transformations

A transformation (or mapping or function) \( T : \mathbb{R}^n \to \mathbb{R}^m \) is a rule that for each \( x \in \mathbb{R}^n \) assigns a vector \( T(x) \in \mathbb{R}^m \), called the image of \( x \).

Matrix multiplication by an \( m \times n \) matrix \( A \) gives a mapping \( \mathbb{R}^n \ni x \to y = Ax \in \mathbb{R}^m \):

\[
\begin{bmatrix}
y_1 \\
y_2 \\
\vdots \\
y_m
\end{bmatrix} =
\begin{bmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
a_{21} & a_{22} & \cdots & a_{2n} \\
\vdots & \vdots & & \vdots \\
a_{m1} & a_{m2} & \cdots & a_{mn}
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
\vdots \\
x_n
\end{bmatrix} =
\begin{bmatrix}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\
\vdots \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n
\end{bmatrix}
\]

or in terms of the rows

\[
y_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\
y_2 = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\
\vdots \\
y_m = a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n
\]

On the other hand one can think of it in terms of the column picture

\[
\begin{bmatrix}
y_1 \\
y_2 \\
\vdots \\
y_m
\end{bmatrix} =
\begin{bmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
a_{21} & a_{22} & \cdots & a_{2n} \\
\vdots & \vdots & & \vdots \\
a_{m1} & a_{m2} & \cdots & a_{mn}
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
\vdots \\
x_n
\end{bmatrix} =
\begin{bmatrix}
a_{11} \\
a_{21} \\
\vdots \\
a_{m1}
\end{bmatrix} +
\begin{bmatrix}
a_{12} \\
a_{22} \\
\vdots \\
a_{m2}
\end{bmatrix} + \cdots +
\begin{bmatrix}
a_{1n} \\
a_{2n} \\
\vdots \\
a_{mn}
\end{bmatrix} \quad (4.1)
\]

A matrix transformation \( T(x) = Ax \) is the simplest type of transformation. It satisfies:

**Th** \( A(x + z) = Ax + Az \) and \( A(\lambda x) = \lambda Ax \), for a scalar \( \lambda \).

**Pf**

\[
\begin{bmatrix}
a_{11} \\
a_{21} \\
\vdots \\
a_{m1}
\end{bmatrix} + \cdots +
\begin{bmatrix}
a_{1n} \\
a_{2n} \\
\vdots \\
a_{mn}
\end{bmatrix} =
\begin{bmatrix}
x_1 \\
x_2 \\
\vdots \\
x_n
\end{bmatrix} + \cdots +
\begin{bmatrix}
x_1 \\
x_2 \\
\vdots \\
x_n
\end{bmatrix}
\]

**Def** A transformation \( T \) is called **linear** if \( T(x + z) = T(x) + T(z) \) and \( T(\lambda x) = \lambda T(x) \).

**Th** If \( T \) is a linear transformation then \( T(x) = Ax \), where

\[
A = \begin{bmatrix}
T(e_1) & T(e_2) & \cdots & T(e_n)
\end{bmatrix}, \quad \text{and} \quad e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \ldots, \quad e_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}
\]

**Pf** We can write

\[
x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1e_1 + x_2e_2 + \cdots + x_ne_n
\]

By linearity and (4.1)

\[
T(x) = T(x_1e_1 + x_2e_2 + \cdots + x_ne_n) = x_1T(e_1) + x_2T(e_2) + \cdots + x_nT(e_n) = Ax.
\]

Note that the book defines linear transformation to be what we call a matrix transformation instead of defining it to be a transformation that has the linearity property.
LINEAR TRANSFORMATIONS DEFINED IN A COORDINATE INVARIANT WAY

The concept of linear transformation can be applied without using specific coordinates. This will be useful in situations where it is difficult to find natural coordinates.

**Ex 1** Let $T$ be the transformation that rotates a vector in the plane 90 degrees counter clockwise. Find an expression for $T$ and deduce that it is linear.

**Sol** This means that:
(i) $\|T(x)\| = \|x\|$, i.e. $T(x)$ has the same length as $x$,
(ii) $T(x) \cdot x = 0$, i.e. $T(x)$ is perpendicular to $x$, and
(iii) $x, T(x)$ are positively oriented, i.e. of the two choices of perpendicular direction we pick the one corresponding to a 90 degree left turn.

If $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ then $T(x) = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix}$ satisfies the three mentioned properties since $\|T(x)\| = \sqrt{(-x_2)^2 + x_1^2} = \|x\|$, $T(x) \cdot x = (-x_2)x_1 + x_1x_2 = 0$, and the orientation is seen from a picture. It follows that $T(x) = Ax$, where $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, which is a linear transformation.

**Ex 2** Suppose that $v_1$ and $v_2$ are two vectors in the plane that are not parallel. Show that there is a linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ such that $T(v_1) = v_1$ and $T(v_2) = 2v_2$.

**Sol** Suppose there is such a transformation that is linear. Since $v_1$ and $v_2$ are not parallel we can find constants $c_1$ and $c_2$, depending on $x$, such that

$$x = c_1v_1 + c_2v_2.$$ 

If we assume that $T$ is linear then

$$T(x) = T(c_1v_1 + c_2v_2) = c_1T(v_1) + c_2T(v_2) = c_1v_1 + 2c_2v_2.$$ 

On the other hand this defines $T$ and it only remains to show that this is linear in $x$.

This will follow once we show below that the constants $c_1$ and $c_2$ are linear functions of $x$.

Now $x$ is clearly a linear functions of $(c_1, c_2)$. In fact

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} v_1 & v_2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}.$$ 

We know that this system can be solved uniquely since $v_1$ and $v_2$ are not parallel.

**Question** Why is that so?

**Def** An inverse $S$ to $T$ is a map such that $S(T(x)) = x$ and $T(S(y)) = y$ for all $x$ and $y$.

**Th** If $T$ is a linear map that has an inverse then the inverse is also a linear map.

**Pf** If $x = Ac$ and if $z = Ab$ and if $S$ is the inverse map so $c = S(x)$ and $b = S(z)$ then

$$x + z = Ac + Ab = A(c + b) = A(S(x) + S(z))$$

and since $S$ is the inverse map we get

$$S(x + z) = S(x) + S(z).$$

**Question** What is the inverse of the map in Ex 1 and what is its matrix?
**Invertibility of a linear transformation**

**Ex 3** Let $T(x) = Ax = y$, where $A = \begin{bmatrix} 1 & 2 \\ 4 & 9 \end{bmatrix}$. If $T$ is invertible find its inverse $S(y) = By$.

**Sol** We want to solve the system

\[
\begin{align*}
x_1 + 2x_2 &= y_1 \\
4x_1 + 9x_2 &= y_2
\end{align*}
\]

Subtracting 4 times the first row from the second gives

\[
\begin{align*}
x_1 + 2x_2 &= y_1 \\
x_2 &= y_2 - 4y_1
\end{align*}
\]

and subtracting 2 times the second from the first gives

\[
\begin{align*}
x_1 &= 9y_1 - 2y_2 \\
x_2 &= y_2 - 4y_1
\end{align*}
\]

i.e. $x = By$, where

\[
B = \begin{bmatrix} 9 & -2 \\ -4 & 1 \end{bmatrix}
\]

**Ex 4** Let $T(x) = Ax = y$, where $A = \begin{bmatrix} 2 & 3 \\ 6 & 9 \end{bmatrix}$. If $T$ is invertible find its inverse $S(y) = By$.

**Sol** We want to solve the system

\[
\begin{align*}
2x_1 + 3x_2 &= y_1 \\
6x_1 + 9x_2 &= y_2
\end{align*}
\]

Subtracting 3 times the first row from the second gives

\[
\begin{align*}
2x_1 + 3x_2 &= y_1 \\
0 &= y_2 - 3y_1
\end{align*}
\]

hence this can not be solve for all $y$ so $T$ is not invertible.

**Question** When is the linear transformation with a $2 \times 2$ matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ invertible? When the columns are not parallel? When the rows are not parallel?
A transformation $T$ is linear transformation, i.e. $T(x+z) = T(x) + T(z)$ and $T(\lambda x) = \lambda T(x)$, if and only if $T$ is matrix transformation, i.e. $T(x) = Ax$, where

$$A = \begin{bmatrix} T(e_1) & T(e_2) & \cdots & T(e_n) \end{bmatrix}, \quad and \quad e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots, e_n = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

What is important with the notion of linearity is that it does not depend on coordinates. This will be useful in situations where it is difficult to find natural coordinates.

If $v_1$ and $v_2$ are two vectors in the plane that are not parallel and $a$ and $b$ are any numbers then there is a linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ such that $T(v_1) = av_1$ and $T(v_2) = bv_2$. In fact since $v_1$ and $v_2$ are not parallel, and $x$ can be written

$$x = c_1 v_1 + c_2 v_2.$$  

If we assume that $T$ is linear then

$$T(x) = T(c_1 v_1 + c_2 v_2) = c_1 T(v_1) + c_2 T(v_2) = c_1 a v_1 + c_2 b v_2.$$  

It remains to show that $c_1$ and $c_2$ depend linearly on $x$ and that follows from the fact:

If $T$ is a linear map that has an inverse then the inverse is also a linear map.

An inverse $S$ to $T$ is a map such that $S(T(x)) = x$ and $T(S(y)) = y$ for all $x$ and $y$.

Non invertibility can be because it is either not onto i.e. there is a $y$ such that $T(x) = y$ has no solution $x$, or there not a unique solution $x$ to $T(x) = y$ for some $y$, i.e. there are $x_1 \neq x_2$ such that $T(x_1) = T(x_2)$.

**Question** When is the linear transformation $\mathbb{R}^2 \ni x \to y \in \mathbb{R}^2$ given by

$$y = Ax, \quad where \quad A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

not invertible? When the columns are parallel? When the rows are parallel?