4. Lecture 4: 2.1 Linear Transformations

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

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

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

or in terms of the rows

$$\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ y_2 &= a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ &\vdots \\ y_m &= a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{aligned}$$

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

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{in} \\ & \vdots & & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} \tag{4.1}$$

A **matrix transformation** $T(\mathbf{x}) = A\mathbf{x}$ is the simplest type of transformation. It satisfies:

Th $A(\mathbf{x}+\mathbf{z}) = A\mathbf{x} + A\mathbf{z}$ and $A(\lambda\mathbf{x}) = \lambda A\mathbf{x}$, for a scalar λ.

Pf

$$\underbrace{(x_1+z_1)\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + \cdots + (x_n+z_n)\begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}}_{A(\mathbf{x}+\mathbf{z})} = \underbrace{x_1\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + \cdots + x_n\begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}}_{A\mathbf{x}} + \underbrace{z_1\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + \cdots + z_n\begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}}_{A\mathbf{z}}$$

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

Th If T is a linear transformation then $T(\mathbf{x}) = A\mathbf{x}$, where

$$A = \begin{bmatrix} | & | & & | \\ T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \\ | & | & & | \end{bmatrix}, \quad \textit{and} \quad \mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \cdots, \mathbf{e}_n = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

Pf We can write

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n$$

By linearity and (4.1)

$$T(\mathbf{x}) = T\big(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n\big) = x_1T(\mathbf{e}_1) + x_2T(\mathbf{e}_2) + \cdots + x_nT(\mathbf{e}_n) = A\mathbf{x}.$$

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) $\|T(\mathbf{x})\| = \|\mathbf{x}\|$, i.e. $T(\mathbf{x})$ has the same length as $\mathbf{x}$,
(ii) $T(\mathbf{x}) \cdot \mathbf{x} = 0$, i.e. $T(\mathbf{x})$ is perpendicular to $\mathbf{x}$, and
(iii) $\mathbf{x}, T(\mathbf{x})$ are positively oriented, i.e. of the two choices of perpendicular direction we pick the one corresponding to a 90 degree left turn.
If $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ then $T(x) = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix}$ satisfies the three mentioned properties since $\|T(\mathbf{x})\| = \sqrt{(-x_2)^2 + x_1^2} = \|\mathbf{x}\|$, $T(\mathbf{x}) \cdot \mathbf{x} = (-x_2)x_1 + x_1 x_2 = 0$, and the orientation is seen from a picture. It follows that $T(\mathbf{x}) = A\mathbf{x}$, where $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, which is a linear transformation.

Ex 2 Suppose that $\mathbf{v}_1$ and $\mathbf{v}_2$ are two vectors in the plane that are not parallel. Show that there is a linear transformation $T : \mathbf{R}^2 \to \mathbf{R}^2$ such that $T(\mathbf{v}_1) = \mathbf{v}_1$ and $T(\mathbf{v}_2) = 2\mathbf{v}_2$.
Sol Suppose there is such a transformation that is linear. Since $\mathbf{v}_1$ and $\mathbf{v}_2$ are not parallel we can find constants c_1 and c_2, depending on $\mathbf{x}$, such that

$$\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2.$$

If we assume that T is linear then

$$T(\mathbf{x}) = T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = c_1 T(\mathbf{v}_1) + c_2 T(\mathbf{v}_2) = c_1\mathbf{v}_1 + 2c_2\mathbf{v}_2.$$

On the other hand this defines T and it only remains to show that this is linear in $\mathbf{x}$. This will follow once we show below that the constants c_1 and c_2 are linear functions of $\mathbf{x}$.

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

$$\underbrace{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}}_{\mathbf{x}} = \underbrace{\begin{bmatrix} | & | \\ \mathbf{v}_1 & \mathbf{v}_2 \\ | & | \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}}_{\mathbf{c}}$$

We know that this system can be solved uniquely since $\mathbf{v}_1$ and $\mathbf{v}_2$ are not parallel.

Question Why is that so?

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

Th If T is a linear map that has an inverse then the inverse is also a linear map.
Pf If $\mathbf{x} = A\mathbf{c}$ and if $\mathbf{z} = A\mathbf{b}$ and if S is the inverse map so $\mathbf{c} = S(\mathbf{x})$ and $\mathbf{b} = S(\mathbf{z})$ then

$$\mathbf{x} + \mathbf{z} = A\mathbf{c} + A\mathbf{b} = A(\mathbf{c} + \mathbf{b}) = A(S(\mathbf{x}) + S(\mathbf{z}))$$

and since S is the inverse map we get

$$S(\mathbf{x} + \mathbf{z}) = S(\mathbf{x}) + S(\mathbf{z}).$$

Question What is the inverse of the map in Ex 1 and what is its matrix?

Invertibility of a linear transformation

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

Sol We want to solve the system

$$\begin{aligned} x_1 + 2x_2 &= y_1 \\ 4x_1 + 9x_2 &= y_2 \end{aligned}$$

Subtracting 4 times the first row from the second gives

$$\begin{aligned} x_1 + 2x_2 &= y_1 \\ x_2 &= y_2 - 4y_1 \end{aligned}$$

and subtracting 2 times the second from the first gives

$$\begin{aligned} x_1 &= 9y_1 - 2y_2 \\ x_2 &= y_2 - 4y_1 \end{aligned}$$

i.e. $\mathbf{x} = B\mathbf{y}$, where

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

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

Sol We want to solve the system

$$\begin{aligned} 2x_1 + 3x_2 &= y_1 \\ 6x_1 + 9x_2 &= y_2 \end{aligned}$$

Subtracting 3 times the first row from the second gives

$$\begin{aligned} 2x_1 + 3x_2 &= y_1 \\ 0 &= y_2 - 3y_1 \end{aligned}$$

hence this can not be solve for all $\mathbf{y}$ so T is not invertible.

Question When is the linear transformation with a 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ invertible? When the columns are not parallel? When the rows are not parallel?

Summary and conceptual questions

A transformation T is linear transformation, i.e. $T(\mathbf{x}+\mathbf{z}) = T(\mathbf{x})+T(\mathbf{z})$ and $T(\lambda\mathbf{x}) = \lambda T(\mathbf{x})$, if and only if T is matrix transformation, i.e. $T(\mathbf{x}) = A\mathbf{x}$, where

$$A = \begin{bmatrix} | & | & & | \\ T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \\ | & | & & | \end{bmatrix}, \quad \textit{and} \quad \mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \cdots, \mathbf{e}_n = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

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 $\mathbf{v}_1$ and $\mathbf{v}_2$ are two vectors in the plane that are not parallel and a and b are any numbers then there is a linear transformation $T : \mathbf{R}^2 \to \mathbf{R}^2$ such that $T(\mathbf{v}_1) = a\mathbf{v}_1$ and $T(\mathbf{v}_2) = b\mathbf{v}_2$. In fact since $\mathbf{v}_1$ and $\mathbf{v}_2$ are not parallel, and $\mathbf{x}$ can be written

$$\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2.$$

If we assume that T is linear then

$$T(\mathbf{x}) = T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2) = c_1a\mathbf{v}_1 + c_2b\mathbf{v}_2.$$

It remains to show that c_1 and c_2 depend linearly on $\mathbf{x}$ 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 $S(T(\mathbf{x})) = \mathbf{x}$ and $T(S(\mathbf{y})) = \mathbf{y}$ for all $\mathbf{x}$ and $\mathbf{y}$.

Non invertibility can be because it is either not **onto** i.e. there is a $\mathbf{y}$ such that $T(\mathbf{x}) = \mathbf{y}$ has no solution $\mathbf{x}$, or there not a unique solution $\mathbf{x}$ to $T(\mathbf{x}) = \mathbf{y}$ for some $\mathbf{y}$, i.e. there are $\mathbf{x}_1 \neq \mathbf{x}_2$ such that $T(\mathbf{x}_1) = T(\mathbf{x}_2)$.

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

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

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